

Early Neural Dynamics of Visual Word Recognition

Yu Li, MSc.

Department of Cognitive Science
ARC Centre of Excellence in Cognition and its Disorders

Faculty of Human Sciences
Macquarie University, Sydney, Australia

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)

November, 2017

Table of Contents

Thesis Abstract	i
Acknowledgments	iii
Author Note	iv
Declaration	v
Abbreviations	vi
Chapter 1: General Introduction	1
Introduction	2
Local brain regions for visual word recognition	4
Inter-regional connections for visual word recognition	9
Neural models of visual word recognition	12
Dynamic causal modeling	16
Time course of visual word recognition	22
Task-dependent neural activity of visual word recognition	25
Outline of this thesis	27
Chapter 2: Early Top-Down Feedback from Frontal to Ventral Occipito-temporal	
Cortex during Visual Word Recognition	30
Abstract	31
Introduction	32
Methods	39
Results	48
Discussion	55
Conclusion	64
Appendix A	65
Appendix B	66
Chapter 3: Task Modulation of Early Top-Down Feedback from Frontal to Ventral	
Occipito-temporal Cortex during Visual Word Recognition	70
Abstract	71

Introduction	72
Methods	79
Results	86
Discussion	96
Conclusion	103
Appendix A	105
Chapter 4: Task Modulation of the Time Course of Visual Word Recognition	109
Abstract	110
Introduction	111
Methods	116
Results	123
Discussion	131
Conclusion	137
Appendix A	139
Chapter 5: General Discussion	140
Summary of Main Findings	142
Neural Models of Visual Word Recognition	145
Task-Dependence of Early Neural Activity	147
Early Stage Functional Integration	149
Appropriateness of Word-Recognition Tasks	151
Limitations and Future Directions	153
Conclusion	158
References	159
Ethics Approval	187

Thesis Abstract

For a skilled reader, visual word recognition can be completed within several hundred milliseconds. There is evidence that the left inferior frontal cortex is activated by visual words in the first 300 ms after stimulus onset; additionally, there is evidence for early top-down feedback from this frontal region to ventral occipito-temporal cortex during visual word recognition. Using magnetoencephalography (MEG), this thesis sought to examine early neural dynamics of visual word recognition by examining the early stage inter-regional connectivity and time course of visual word recognition.

In Chapter 1, I review studies of the neural correlates and relevant neural models of visual word recognition; in particular, two models on the ventral occipito-temporal cortex in visual word recognition with contrasting views on the role of top-down feedback are examined. I then introduce dynamic causal modeling (DCM), a crucial neuroimaging method for examining directional influences of one brain area to another. Next I review the neuroimaging studies of visual word recognition focusing on its time course and also highlight the importance of examining early brain activity of visual word recognition. Finally, I propose the research questions to be addressed in this thesis: What is the nature of early top-down feedback from frontal to ventral occipito-temporal cortex during visual word recognition? How task goals modulate the early top-down feedback? How task goals modulate the time course of visual word recognition? These questions are examined in three empirical chapters.

Using a semantic categorisation task, Chapter 2 examines the nature of top-down feedback from the left inferior frontal gyrus (LIFG) to the left ventral occipito-temporal cortex (LvOT) during the first 200 ms visual word recognition. The results revealed that the LIFG-to-LvOT connection was stronger for real words than for pseudowords, and stronger for false fonts than for consonant strings in both 1-150 ms and 1-200 ms time-windows, indicating that both lexical-semantic and surface letter information influence early top-down feedback. Furthermore, the LIFG-to-LvOT connection was stronger for pseudowords than for consonant strings in the 1-200 ms time-window, indicating that compared with lexical-semantic and surface letter information, the influences of phonological information occur later.

By comparing a non-linguistic visual discrimination task (is it a hash string?) with the semantic categorisation task (is it an animal word?) used in Chapter 2, Chapter 3 examines how task goals modulate the early LIFG-to-LvOT feedback during the first 200 ms visual word

recognition. The results revealed that the LIFG-to-LvOT connection was stronger for real words than for pseudowords in both 1-150 ms and 1-200 ms time-windows in the discrimination task, and this was similar to that in the semantic task, indicating that the influences of lexical-semantic information on this feedback are independent of task goals. However, in the discrimination task the LIFG-to-LvOT connection was stronger for consonant strings than for pseudowords in the 1-200 ms time-window and was stronger for consonant strings than for false fonts in the 1-150 ms time-window, and this was opposite to that in the semantic task, indicating that the influences of letter and phonological information on the early feedback are dependent on task goals.

Using the data from the previous two chapters, Chapter 4 examines how task goals modulate the on-line time course of visual word recognition by examining the first 500 ms time course data during the linguistic semantic task and the non-linguistic discrimination task. A behavioural version of each MEG experiment with one task is also reported on. The MEG and behavioural results established that an emphasis on high-level linguistic information in a linguistic task sensitise early neural responses to linguistic properties, whereas an emphasis on low-level visual feature detection in a non-linguistic task sensitise early neural responses to physical appearance.

This thesis has established that lexical-semantic, phonological, and letter properties all can trigger early top-down feedback during visual word recognition, but occur differently across time. Task goals with different emphases strongly bias early top-down feedback and also modulate the time course of visual word recognition.

Acknowledgments

Pursuing a higher research degree in Australia is an excellent decision in my life. My PhD journey is beautiful in that over the past three years I met kind and intelligent scholars, had different experiences, and more importantly learnt new knowledge. This thesis would not have been possible without the support of many others.

First and foremost, I would like to express my gratitude to my fabulous supervisors, Distinguished Prof Anne Castles, A/P Sachiko Kinoshita, and A/P Paul Sowman for their excellent supervision and continuous support over the past three years. Thank you all for sharing your knowledge, providing technical support, and encouraging me. You all are a constant source of wisdom and inspiration. I have learnt so much from you not only about research but also life wisdom.

Thanks to Erika Contini, Valerie Pu, Yanan Sun, Erin Martin, and Elisabeth Magdas for helping me collect MEG data, and two MQ students Claris Teng and Alol Elba for coding behavioural results. Thank all subjects for participating in the experiments that make this thesis possible. Thanks to Dr Andrew Etchell for providing help with the MEG data analysis. Many thanks to Dr Zoe Woodhead for sharing a script and Dr Marta Garrido for answering some technical questions. Thanks to Dr Jane Warren for sharing “Carian” font used in this thesis.

Many thanks to the CCD administrative team, especially thanks to Lesley McKnight, Teri Roberts, Katie Webb, Anne Van Uden, Marcus Ockenden, and Craig Richardson for their support. I would like to thank CCD people for bringing new ideas, chatting with me, forming a beautiful and inspiring research atmosphere.

I would like to thank Prof Bruce McCandliss for hosting me at Stanford University and Dr Jason Yeatman for hosting me at Washington University in Seattle. Thank you for your valuable suggestions on my PhD work.

I would like to thank Dr Arthur Adams and Ross Carruthers for encouragement and help in the past three years. Thanks to my great parents for their endless support from north hemisphere. A huge thank to my fiancée Sabrina Yanan Jiang for her everlasting love.

Author Note

This thesis was presented in the form of thesis by publication. Each empirical chapter in this thesis presents a stand-alone research article. Each empirical chapter will be submitted to an academic journal. References are given at the end of this thesis.

Declaration

I, Yu Li, certify that the work in this thesis entitled “Early neural dynamics of visual word recognition” has not previously been submitted for a degree and submitted to any university other than Macquarie University.

I also certify that this thesis presents an original piece of research and that it has been written by me. Any assistance I have received has been appropriately acknowledged. All information sources and literature have been cited.

The research presented in this thesis was approved by the Macquarie University Human Research Ethics Committee (reference number: 5201500186). The approval letter is attached.

Signed

A handwritten signature in black ink, appearing to be the Chinese characters '李雨' (Li Yu), written in a cursive style.

Yu Li

November 2017

Abbreviations

ALE = Activation Likelihood Estimation

BA = Brodmann area

BMA = Bayesian Modeling Averaging

BMS = Bayesian Modeling Selection

BOLD = Blood-Oxygen-Level-Dependent

CS = consonant strings

DCM = Dynamic Causal Modelling

EEG = Electroencephalography

ERF = event-related field

FDR = False-discovery Rate

FF = false fonts

fMRI = Functional Magnetic Resonance

Imaging

KBIT = Kaufman Brief Intelligence Test

LCD = Local Combination Detector

LGN = lateral geniculate nucleus

LIFG = left inferior frontal gyrus

LpSTC = left posterior superior temporal

cortex

LvOT = left ventral occipito-temporal

cortex

MEG = Magnetoencephalography

MNI = Montreal Neurological Institute

MRI = Magnetic Resonance Imaging

OCC = primary occipital cortex

PET = Positron Emission Tomography

PW = pseudowords

RMS = Root Mean Square

RW = real words

TOWRE = Test of Word Reading

Efficiency

VWFA = visual word form area

Chapter 1

General Introduction

Introduction

Language is arguably the most important tool in the history of human evolution. Spoken language was one of the primary forms of communication for exchanging ideas, but since the invention of writing systems, written language has been added to the toolbox of human communication. The first writing system dates back to only about 5400 years ago (Daniels, 1996), thus literacy is not an innate skill, but has to be acquired through experience. As one of the most important elements in modern education, literacy is a prerequisite for acquiring and sharing knowledge, and achieving personal and career success (Elliott & Grigorenko, 2014). When a child starts to learn a writing system, visual symbols of the writing system are linked to corresponding speech sounds and meanings. With intense training and experience, the links between these bodies of knowledge become stronger and stronger so that when a visual word is displayed, one can rapidly retrieve its meaning and sound. Skilled visual word recognition involves different processing stages from low-level visual feature detection to orthographic form analysis and high-level phonological and lexical-semantic processing (e.g., Rastle, 2015). Put simply, visual word recognition is a journey from visual features to meaning (Balota, 1994).

With the development of neuroimaging technology in the past 20 years, researchers have gained insights into the neural basis of visual word recognition. Functional magnetic resonance imaging (fMRI) studies have revealed that the journey from visual features to meaning involves distributed local brain regions including primary visual cortex, ventral occipito-temporal cortex, and spoken language regions such as the left inferior frontal gyrus (Taylor, Rastles & Davis, 2013). These explorations have also demonstrated that this journey involves neural interactions between these relevant regions, especially between the posterior ventral occipito-temporal cortex and spoken language regions (Price, 2012; Taylor et al., 2013; Martin, Schurz, Kronbichler, & Richlan, 2015). Indices of these neural interactions include

functional connectivity measuring neural couplings in time-series between brain regions and effective connectivity measuring directional influences of one region on another (Friston, 2011). There is accumulating evidence that these local brain activities and inter-regional neural interactions are modulated by task goals (Pugh et al., 2000; McNorgan, Chabal, O’Young, Lukic, & Booth, 2015; Bitan et al., 2005, 2006; Heim et al., 2009).

There is mounting evidence that the cognitive processes underpinning visual word recognition unfold rapidly (Serenio & Rayner, 2003; Carreiras, Armstrong, Perea, & Frost, 2014). For skilled readers, a visual word can be recognised within several hundred milliseconds, by which time lexical-semantic and phonological information has already been accessed (Keuleers, Diependaele, & Brysbaert, 2010; Keuleers, Lacey, Rastle, & Brysbaert, 2012; Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson, 2006; Carreiras et al., 2014; Hauk, 2016). Electrophysiological studies have shown that the left inferior frontal gyrus is robustly involved in the early stage visual word recognition in that this region has already been activated within 200 ms after visual word onset (e.g., Pammer et al., 2004; Cornelissen et al., 2009). A recent effective connectivity study further observed that this frontal region sends early feedback to the left ventral occipito-temporal cortex within 200 ms after word onset (Woodhead et al., 2014). These findings have widened and updated knowledge about the early neural dynamics of visual word recognition.

However, the nature of early top-down feedback from the left inferior frontal gyrus to the left ventral occipito-temporal cortex – whether it represents semantic, phonological or orthographic processing - remains largely unknown. Previous fMRI studies have also indicated that the inter-regional connectivity for visual word recognition is dependent on task goals (Bitan et al., 2005, 2006). This raises an outstanding question - how task goals modulate this early top-down feedback. The evidence for rapidly unfolding cognitive processes during visual word recognition also leads to the question of how task goals modulate the time course of visual

word recognition - the journey from visual features to meaning. The aim of this thesis was to address these outstanding questions regarding the early neural dynamics of visual word recognition in three separate empirical studies, by using magnetoencephalography.

In this General Introduction, I will introduce relevant background on the neural basis of visual word recognition by reviewing extant neuroimaging studies, neural models, and important techniques. I first provide a literature review of the neural correlates of visual word recognition including local brain regions, inter-regional connectivity, and relevant neural models, which will be used in Chapters 2 and 3. I then provide a detailed description of dynamic causal modeling measuring inter-regional causal influences, which will be used as the key method in Chapters 2 and 3. I then systematically review the necessity of using electrophysiological signals with high temporal resolution in the examination of the early neural dynamics, and existing findings, which are highly relevant to all empirical studies in this thesis. Subsequently, I provide a review of task-dependent neural activity and automaticity of visual word recognition, which will be relevant to Chapters 3 and 4. Finally, I provide a brief overview of each chapter including the research focus and key method to be used.

Local Brain Regions for Visual Word Recognition

Neural studies have revealed that multiple local brain regions are involved in visual word recognition. To date, the neural bases of visual word recognition have been informed by neuropsychological data from patients with brain-damage and neuroimaging techniques including positron emission tomography (PET), magnetoencephalography (MEG), electroencephalogram (EEG), and functional magnetic resonance imaging (fMRI).

Neuropsychological Studies

Early findings regarding the brain areas playing a role in language were contributed by studies of patients with brain injuries. Paul Broca found that an area of the left frontal lobe was damaged in an aphasic patient who was without speech (Broca, 1861). Because he was unable to produce words but could still understand spoken words, Broca argued that the frontal area might have a function specifically for speech production. This area is now typically defined as the pars opercularis and pars triangularis of the inferior frontal gyrus (LIFG). Karl Wernicke later observed that damage to the left posterior superior temporal gyrus led to language comprehension deficits (Wernicke, 1874). This area is typically defined as the posterior part of the left superior temporal cortex (LpSTC). The damage to Broca's area and Wernicke's area thus result in speech production and language comprehension respectively. Given that visual word recognition is a process of linking visual symbols with phonology and meaning, these areas play a very crucial role in visual word recognition.

What brain areas are specifically engaged during visual word recognition? Joseph Dejerine described a patient in which a lesion of the left occipito-temporal cortex (LvOT) caused selective loss of the ability to read letters and words, but the patient's visual ability, writing ability and auditory language comprehension were spared (Dejerine, 1892). It was interpreted that this area is selectively responsible for recognising the orthographic form of words and is a crucial area linking the visual system with language areas (e.g., LIFG and LpSTC). Damage to this area results in disconnections from visual inputs to language areas for auditory form processing and articulation. Supporting the neuropsychological evidence, Mani et al. (2008) found that electrical stimulation to sites within the basal occipito-temporal cortex resulted in selective reading difficulty in three participants. Auditory comprehension and writing remained intact and there was no impairment in object naming, confirming the crucial

role of the LvOT in visual word recognition. The important role of the LvOT in reading and visual word recognition has been supported by investigations with modern neuroimaging techniques, as described below.

Modern neuroimaging Studies

Evidence from patient studies has suggested that the LIFG, LpSTC and LvOT are involved in different facets of language processing. Specifically, LIFG is mainly responsible for speech production or articulation, LpSTC for speech comprehension and LvOT for visual word processing. These interpretations have been verified by modern brain imaging techniques which can be used to measure brain activity noninvasively. The most widely-used neuroimaging techniques include PET, fMRI and MEG/EEG. Functional MRI measures brain activity by recording a Blood-Oxygen-Level-Dependent (BOLD) signal; that is, detecting blood flow changes in the brain (Ogawa, Lee, Kay, & Tank, 1990; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001). It relies on the fact that cerebral blood flow and neuronal activation are coupled so that the BOLD signal provides a localisable marker of neuronal activity (Huettel, Song & McCarthy, 2009). Thus, the underlying neuronal activity within an area can be inferred based on its BOLD signal. With the excellent spatial resolution, fMRI helps to uncover the potential local brain areas which respond to visual words. Studies using reading-related tasks, such as lexical decision and reading aloud, have shown that widely-distributed areas in the brain are recruited during visual word recognition or reading aloud (e.g., Turkeltaub, Eden, Jone, & Zeffiro, 2002; Bolger, Perfetti, & Schneider, 2005; Richlan, Kronbichler, & Wimmer, 2009, 2011; Houdé, Rossi, Lubin, & Joliot, 2010; Price, 2012; Taylor et al., 2013; Cattinelli, Borghese, Gallucci, & Paulesu, 2013; Zhu, Nie, Chang, Gao, & Niu, 2014; McNorgan et al., 2015; Martin et al., 2015; Zhao, Fan, Liu, Wang, & Yang, 2017).

The activation likelihood estimation (ALE) meta-analysis, originally developed by Turkeltaub et al. (2002) and further modified by Laird et al. (2005) and Eickhoff et al. (2009, 2012), is a quantitative method that considers a number of studies to determine local brain regions consistently seen in neuroimaging research. As a coordinate-based meta-analysis method, ALE considers reported local brain activations as spatial probability distributions centred at given coordinates. ALE maps are then generated by computing the convergence of these activation probabilities for each voxel. ALE results are then assessed by a permutation procedure to against a null-distribution of random spatial activation across studies. Several meta-analyses of fMRI and PET studies of visual word recognition have been conducted to identify key local brain regions (e.g., Turkeltaub et al., 2002; Bolger et al., 2005; Houde et al., 2010; Cattinelli et al., 2013; Martin et al., 2015; McNorgan et al., 2015). Turkeltaub et al. (2002) conducted a meta-analysis of 11 PET studies with single word reading tasks and identified the bilateral motor, superior temporal cortex, supplementary motor area and LvOT as being key areas. Houde et al. (2010) specifically examined the brain areas for reading behaviour in children and found that the LvOT, LpSTC, and the bilateral IFG are implicated. Martin et al. (2015) further found brain activation common to both children and adults in the LvOT, LIFG, and posterior parietal regions, and revealed higher convergence in studies with children in the left superior temporal and bilateral supplementary motor regions, but higher convergence in studies with adults in the bilateral vOT and left dorsal precentral regions. These observed developmental changes and differences between children and adults probably reflect reading experience and its roles in brain plasticity.

There is also evidence that these reading-related areas are aberrant, hyperactive or hypoactive in people with dyslexia. For example, Maisog, Einbinder, Flowers, Turkeltaub, & Eden (2008) found that typical readers show greater activation than dyslexics in the left extrastriate areas, inferior parietal cortex, superior temporal gyrus, thalamus, and left inferior

frontal gyrus; and hyperactivity associated with dyslexia was found in the right thalamus and anterior insula. Among these findings, the most robust was in relation to the LvOT, where reduced activation was associated with dyslexia. Richlan et al. (2009) found that dyslexic people had underactivation in the left inferior parietal, superior temporal regions, as well as in the LIFG. Richlan et al. (2011) further found a common underactivation in LvOT in both children and adults with dyslexia. Both Richlan et al. (2009, 2011) and Maisog et al. (2008) thus converge in finding that the LvOT is aberrant in dyslexics.

Although there is some debate surrounding differences in the neural underpinnings of different writing systems (e.g., Bolger et al., 2005; Zhu et al., 2014; Wang et al., 2015; Rueckl et al., 2015), all of these studies and other review studies (e.g., Price, 2012) tend to agree in concluding that visual word recognition consistently recruits three broad areas: the LIFG, LpSTC, and LvOT. Many studies have also shown that subcortical structures, such as the putamen, thalamus and caudate, are also involved in normal visual word recognition and are aberrant in individuals with reading impairments (Shaywitz et al., 2002; Booth, Bebko, Burman, & Bitan, 2007; Maisog et al., 2008; Richlan et al., 2009, 2011; Seghier & Price, 2009; Oberhuber et al., 2013; Hancock, Richlan & Hoefft, 2017).

Therefore, local regions distributed in the brain involved in visual word recognition mainly include the left ventral occipito-temporal cortex (LvOT), the left posterior temporal cortex (LpSTC), the left inferior frontal gyrus (LIFG) and subcortical areas (Pugh et al., 2000; Price & Mechelli, 2005; Taylor et al., 2013). These regions have different functional roles in visual word recognition. The extant neuroimaging evidence has formed a general functional anatomical map in which activation in the LvOT reflects orthographic form analysis, activation in the LpSTC reflects auditory form processing including orthography-to-phonology conversion, and activation in the LIFG reflects phonological output manipulation and lexico-

semantic processing (Taylor et al., 2013; Carreiras et al., 2014). See “neural models of visual word recognition” for details.

Inter-Regional Connections for Visual Word Recognition

The neuroimaging literature reviewed so far has indicated that visual word recognition involves different regions for distinct functions. Visual word recognition involves linking visual symbols to spoken language knowledge such as phonological and lexical-semantic information. Thus, these different local brain regions have to be connected so that visual word recognition can be completed. Indeed, past studies have revealed that these distributed local brain regions are not only activated by related word recognition tasks, but also are functionally or structurally connected to each other.

Local brain regions involved in visual word recognition are connected via functional couplings. Functional connectivity is defined as the statistical association or dependency among two or more anatomically distinct time-series (Friston, 1994, 2011). Much work has contributed to the understanding of functional connectivity between reading-related areas. Some have examined the inter-regional functional connectivity in the brain when a reading task is performed (task-based, e.g., Hampson et al., 2006; van der Mark et al., 2011; Finn et al., 2014; Boets et al., 2013; Schurz et al., 2014). Others have examined the inter-regional functional connectivity in the brain when no specific task is performed (task-free, e.g., Koyama et al., 2011; Vogel, Miezin, Petersen, & Schlaggar, 2011; Wang, Han, He, Liu, & Bi, 2012; Zhang et al., 2014; Li et al., 2017; Alcauter et al., 2017). Notably, Koyama et al (2011), Wang et al. (2012) and Li et al. (2017) found that the LvOT-LIFG and LvOT-LpSTC connectivity could predict reading competence in normal child and adult readers. Li et al. (2017) found that compared with children, the LvOT-LIFG and LvOT-LpSTC connectivity was stronger in

young adults, indicating stronger relationships between the orthographic form region LvOT and the spoken language network in adults. These developmental changes in brain connectivity may also explain the differences in language network activation between children and adults (Martin et al., 2015). van der Mark et al. (2011) and Finn et al. (2013) found that, compared with normal readers, the LvOT-LIFG and LvOT-LpSTC connectivity was disrupted or reduced in people with dyslexia. These studies collectively indicate that functional couplings between reading-related areas during a task or at rest can index variations in reading ability; and disrupted connections can be used as a neural marker of dyslexia or poor reading. Specifically, the LvOT as a region for visual word processing is functionally coupled with other spoken language areas.

In contrast to functional connectivity, effective connectivity refers to the causal architecture of coupled or distributed dynamical systems (Friston, 2011). In other words, effective connectivity examines inter-regional directional connections. A widely-used method used to investigate effective connectivity is dynamic causal modeling (DCM; Friston, Harrison & Penny, 2003; see “Dynamic Causal Modeling” for a detailed description). DCM aims to model couplings among the hidden states generating observations and is specifically concerned with directed causal interactions (Friston et al., 2003; Friston, 2011). DCM has been widely used in fMRI studies of visual word processing (Bitan et al., 2005, 2006, 2007; Cao, Bitan, & Booth, 2008; Booth, Mehdiratta, Burman, & Bitan, 2008; Heim et al., 2009; Liu et al., 2010; Schurz et al., 2014; Xu, Wang, Chen, Fox, & Tan, 2015; Morken, Helland, Hugdahl, & Specht, 2017; Perrone-Bertolotti, Kauffmann, Pichat, Vidal, & Baciú, 2017). These studies collectively provide strong evidence that the LvOT plays a crucial role in visual word recognition by inputting visual sensory information from occipital cortex and receiving linguistic information from language areas such as the LIFG. For example, Bitan et al. (2005, 2006) found significant intrinsic connections from LvOT to LIFG and from LIFG to LvOT in both children and adults;

Perrone-Bertolotti et al. (2017) found that the strength of the connection from the LvOT to the ventral LIFG specifically increased during a semantic task; and both Heim et al. (2009) and Schurz et al. (2004) observed causal influences between the LIFG and the LvOT. Functional and effective connectivity used in reading studies reveal how information in the brain flows during visual word recognition.

The information flow in the reading network is supported by structural connectivity; that is, white matter pathways. Investigations of the language network have found that between the LIFG for speech production and the LpSTC for language comprehension is a white matter pathway called arcuate fasciculus (Dejerine, 1895; Catani, Jones & Ffytche, 2005). Further studies have even found that this dorsal white matter pathway exists in newborn infants (e.g., Perani et al., 2011). Also observed is a ventral white matter pathway connecting the LpSTC to the LIFG (e.g., Perani et al., 2011; Brauer, Anwander, Perani, & Friederici, 2013). In relation to visual word recognition, white matter pathways have been identified connecting the arcuate fasciculus to the LvOT and LpSTC/LIFG, the inferior fronto-occipital fasciculus to the LvOT and the LIFG, the inferior longitudinal fasciculus to the occipital cortex to anterior temporal lobe, and the vertical occipital fasciculus to the parietal cortex, inferior parietal lobule and the LvOT (Ben-Shachar, Dougherty, & Wandell, 2007; Yeatman, Rauschecker, & Wandell, 2013; Vandermosten, Boets, Wouters, & Ghesquière, 2012). Increasing evidence from behaviour-brain correlations is that the integrity of these white matter pathways is correlated with reading-related skills in children and adults (Klingberg et al., 2000; Yeatman et al., 2011; Yeatman, Dougherty, Ben-Shachar, & Wandell, 2012; Gullick & Booth, 2015; Saygin et al., 2013; see reviews by Ben-Shachar et al., 2007, Vandermosten et al., 2012; Wandell, Rauschecker, & Yeatman, 2012).

In summary, the LvOT, LpSTC, LIFG, and subcortical structures are local brain regions distributed in the brain that are not only consistently elicited by visual words, but also

functionally interact with each other and are structurally connected via white matter pathways. These single local regions and the connections between them play an important role in visual word recognition and its atypical forms such as dyslexia. Neural models have been developed to interpret these findings in relation to reading and visual word recognition under a unified framework.

Neural Models of Visual Word Recognition

Several neural models have been developed to explain what brain regions are involved in visual word recognition and how information flows between these regions (Dejerine, 1981; Geschwind, 1965; Pugh et al., 2000; Price, 2000; Price & Mechelli, 2005; Dehaene, Cohen, Sigman, & Vinckier, 2005; Dehaene & Cohen, 2011; Price & Devlin, 2011). Although classic models such as those of Dejerine (1981) and Geschwind (1965) made an important contribution to knowledge, the focus here is on recent neural models of visual word recognition that have been informed by the latest knowledge of key reading regions in the brain, functional-anatomic mappings, and the neural dynamics of reading.

Recent Neural Models

Pugh et al. (2000) have proposed a comprehensive neural model of reading. In this model, reading involves three different parts: a ventral posterior system in the ventral occipito-temporal cortex, a dorsal posterior system in the temporo-parietal cortex including angular and supramarginal gyri, and an anterior system based in the inferior frontal cortex. The ventral posterior system is proposed to be associated with memory-based word identification and to be responsible for visual word form analysis. The dorsal posterior system is associated with rule-based functional integration of orthographic, phonological and lexical-semantic

information. The anterior system is associated with fine-grained articulatory recoding. In typically developing readers, the dorsal circuit is proposed to be an early-developing word identification system while the ventral circuit is a fast and late-developing system. With the increasing findings on the reading brain, reading models have been further refined (Price, 2000; Price & Mechelli, 2005). The basic anatomical frame of the new and the old models is similar. The novel parts of new models are that each system is divided into subregions with different functional roles, the systems are connected via white matter pathways, and information is allowed to flow between systems in both feedback and feedforward manners.

These models developed by Pugh et al. (2005) and Price and Colleagues (2000, 2005) help to understand with more precision how visual word recognition takes place in the brain, especially within local brain regions associated with visual words. However, these modern models were mainly built on evidence of fMRI and PET neuroimaging techniques with good spatial resolution. However, the neural models above are limited in interpreting the rapid neural dynamics of visual word recognition because cognitive processes taking place after a visual word is presented unfold rapidly (Serenio & Rayner, 2003; Carreiras et al., 2014). The field calls for the application of neuroimaging with high temporal resolution to the establishment of reading models.

Neural Models of Processing in the Left Ventral Occipito-Temporal Cortex

Recent years have witnessed striking progress in the understanding of the ventral posterior system, the left ventral occipito-temporal cortex (LvOT), in visual word processing. Here, I introduce two models of this region, the Local Combination Detector model (Dehaene et al., 2005) and the Interactive Account (Price & Devlin, 2011).

Dehaene and colleagues (2005) built the Local Combination Detector (LCD) model by taking advantage of the large amount of knowledge about the organisation of the ventral

occipito-temporal stream for object recognition in humans and macaques (Riesenhuber & Poggio, 1999; Rolls, 2000; Malach et al., 2002), especially the crucial properties relevant to reading, such as hierarchical organisation. Receptive field size, visual complexity, location-invariance and coded units increase from lateral geniculate nucleus to V1, V2, V4, V8, and then to left occipito-temporal sulcus. Low-level sensitivity takes place in the low part of ventral stream while high-level sensitivity takes place in the high part. When a visual word comes in, different areas along the visual ventral stream are tuned to different letter, letter combination and word features in a hierarchical manner. Simple features, such as local contrasts in lateral geniculate nucleus and orientation in V1, are encoded before the later stage V2 encodes letter fragments and V4 encodes letter shapes. After that, abstract letters are encoded in V8. All of these processes occur bilaterally. Then, local bigrams are encoded in the posterior part of the left occipito-temporal sulcus and small words and substrings such as morphemes are encoded in the anterior part. All processes at different stages proceed in a feedforward manner, i.e., from posterior to anterior parts of visual cortex.

According to this model, it is predicted that the anterior part of the left ventral stream will be more sensitive to words than word-like stimuli and nonwords, and this has been supported by the results of an fMRI study (Vinckier et al., 2007). Because of the preferential response of this area to visual words and its location reproducibility, Dehaene and colleagues label it the visual word form area (VWFA; McCandliss, Cohen & Dehaene, 2003; Dehaene & Cohen, 2011). The LCD model has many advantages. It explicitly states what types of information related to visual words are encoded in what parts of the brain. It also states that the late stages of visual word processing such as local bigram and substring encoding are processed in the left hemisphere - probably due to the left-lateralisation of language network - and thus provides an account of how the VWFA can send information to spoken language networks for phonological and semantic processing. However, this model also has some limitations. It does

not include feedback connections within the visual cortex or between the visual cortex and spoken language areas. The LCD model posits that the VWFA is an area specific for written words that computes strictly visual and abstract pre-lexical orthographic representations in a primarily feed-forward manner. In contrast to the LCD model, the Interactive Account of LvOT has a very different approach.

Price and Devlin (2011) proposed the Interactive Account to interpret activity in the LvOT during reading. The premise of this theory is that perception involves recurrent or reciprocal communications between sensory cortices and high-order areas via a hierarchical forward and backward connections (Friston, 2010). The function of a region is thus proposed to depend on its integration of bottom-up sensory inputs via forward connections as well as top-down predictions via backward connections. These predictions depend on prior experience that can be used to resolve uncertainty and ambiguity of sensory inputs.

According to the Interactive Account, during reading, visual words are processed as sensory inputs via bottom-up connections and receive predictions about phonological and semantic properties via top-down connections, which takes place in the vOT. Therefore, the vOT is an interface linking low-level visual sensory inputs and high-level phonological and semantic regions. The vOT itself is not proposed to be specific to written words *per se*. When there is no prior knowledge of the associations between orthography and phonology, the response of vOT to a word will be low because there is no top-down prediction sent by phonological regions. If the knowledge is learnt, phonological regions send predictions to the vOT when a visual word or word-like stimulus is presented. The intimate association between visual inputs and linguistic representations occurs automatically and is modulated by attention and task demands. Therefore, the interpretation of the vOT responses to words depends on the stimulus, experience, and task context. An important feature is that the responses of neurons across the vOT to basic shape information about visual words are sufficient to partly activate

neurons encoding phonology and semantics in high-order regions, which can in turn provide recurrent inputs to the vOT until the top-down predictions and bottom-up inputs are maximally consistent. Thus, top-down predictions are refined during the integration of top-down and bottom-up information. The main difference between the LCD model and the Interactive Account about the role of LvOT in reading and visual word recognition is that the former proposes the LvOT to be a region specific for word form representations while the latter argue for the LvOT as an interface between bottom-up visual inputs and top-down predictions from high-order phonological and semantic regions. These two models of the LvOT provide different perspectives on how reading works in the brain.

The Interactive Account of reading calls for the use of effective connectivity in examining inter-regional forward and backward connections in the reading brain, specifically measures of how the LvOT interacts with spoken language regions. A popular means of doing this is via dynamic causal modeling (DCM). Because cognitive processes in visual word recognition are fast-acting (Serenio & Rayner, 2003; Carreiras et al., 2014), the application of DCM to M/EEG data with high-temporal resolution can specifically contribute to the examination of early interactions between reading-related regions.

Dynamic Causal Modeling

DCM has become one of the commonly-used methods to examine forward and backward connections in neural networks (Friston, Harrison & Penny, 2003; Friston, 2011; Daunizeau, David & Stephan, 2011). DCM was originally introduced for fMRI time-series data (Friston et al., 2003) and then later for electromagnetic data (David et al., 2006; Kiebel et al., 2009; Chen, Kiebel, & Friston, 2008; Penny, Litvak, Fuentemilla, Duzel, & Friston, 2009).

Central Idea of DCM

DCM adopts the concept of effective connectivity which refers to the influences of one neuronal system on another. The central idea of DCM is that neuronal activity flows through brain networks in a deterministic, nonlinear, and dynamical system that incorporates inputs, hidden states, and outputs. Inputs including experimental manipulations (e.g., stimulus presentation) and/or experimental factors (e.g., stimulus attributes or task demands) evoke responses in the brain. Hidden state variables include neuronal activities and other neurophysiological variables that form the outputs. Outputs are observable measured responses. A neuronal model with inputs and outputs is augmented with a forward or observational model describing the mapping from hidden neuronal activity to observed responses. Neuronal and observational models together form a full generative model. The term “generative” means that “a DCM can be regarded as a prescription of how the observed data were generated” (Kiebel, Klöppel, Weiskopf, & Friston, 2007, p. 1487).

With observed neuronal activity or prior knowledge of neuronal activity elicited by tasks/stimuli, one can establish a neuronal model with interactions among cortical regions. The neuronal model’s parameters, namely effective connectivity within or between regions, are then estimated from observed responses and the way in which these parameters are influenced by experimental manipulations or factors. Using marginal likelihood or evidence over conditional density of the model parameters, one can estimate the probability of the observed responses for a particular model. The outputs of DCM are the evidence for different models with different cortical regions and/or connections, especially those describing couplings among brain regions. Using a Bayesian model selection procedure, one can determine which model is the best model or which models with common features form the best model family. This enables one to test hypotheses about how remote brain regions communicate, especially how

these communications are influenced by experimental manipulations. DCM is a model-based method to estimate effective connectivity because it relies on prior knowledge of neuronal activity, especially key nodes and connections included in a model. Thus, one can select a model that best explains the measured responses from a finite selection of models, but cannot determine that the selected model is a definitive representation of neuronal activity (Kiebel et al., 2007). See Friston et al. (2003) and Daunizeau et al. (2011) for a detailed description.

DCM for Evoked Responses

Spatiotemporal models are designed to examine the architecture of underlying neuronal dynamics and to make inferences about key neuronal parameters in M/EEG data (Kiebel et al., 2009). DCM has been used to analyse evoked responses, induced responses and steady-state responses. The aim of the current thesis was to examine the early neural dynamics estimated by average evoked MEG responses. Therefore, the technical focus here was on the application of DCM for evoked responses in MEG data. Note that DCM is not limited to the neural mass model described below, and DCM is very flexible so that one can design specific models based on specific research aims or hypotheses (Kiebel et al., 2009; Litvak et al., 2011).

MEG signals. The MEG technique is completely non-invasive. It measures the magnetic fields generated by synchronized neuronal currents in the brain. As the magnetic fields are very weak, MEG measurements are made in a magnetically shielded room to reduce magnetic noise from other sources. Unlike BOLD signals recording changes in blood flow in the brain, MEG records neuronal activity directly, the measurements are thus thought to be directly related to brain functions. This technique has excellent temporal resolution (milliseconds) and good spatial resolution (millimetres). In the cognitive neuroscience field, the MEG technique is being increasingly used to examine brain functions including language and attention.

Neural mass model. M/EEG data can be treated as the response of a dynamic input-output system to experimental manipulations or factors. Sensory inputs (stimuli) are processed by a brain network consisting of local neuronal sources and interactions between them. In the DCM for evoked responses, a neural mass model (Jansen & Rit, 1995) is used to describe the neuronal dynamics of each single source. In the model, each source comprises three neuronal subpopulations, pyramidal cells in the infragranular layer, spiny stellate cells in the granular layer, and inhibitory interneurons in the supragranular layer (see Figure 1 in Kiebel et al., 2009). The average post-membrane potentials and mean firing rates of these three neuronal subpopulations are used to describe each single source. Mean firing rates from other sources arrive via directed forward (or bottom-up), backward (or top-down), and lateral connections. Forward connections originate in the infragranular layer and terminate in the granular layer; backward connections connect infragranular to granular layers; lateral connections originate in infragranular layer and end in all layers (see Figure 1 in Kiebel et al., 2009). All these extrinsic cortico-cortical connections are excitatory and are mediated through the axons of pyramidal cells. Three different neuronal subpopulations within a single source are connected by intrinsic connections. These extrinsic cortico-cortical and intrinsic connections are used to estimate bottom-up and top-down brain connectivity (see Jansen & Rit (1995) for the details of the neural mass model).

Neuronal sources. The depolarisation of pyramidal cell populations is assumed to be the origin of M/EEG responses. These sensor-level responses are expressed through a lead-field where each source corresponds to an equivalent current dipole (Kiebel, David & Friston, 2006). A lead-field is an electric current field describing the relationship between sources and sensors; a dipole describes the strength and direction of the current flow of the summed activity within a specific source. This is used to model observed activity in the sensors (see Figure 1 in Kiebel et al., 2009). The location of each source can be obtained from source reconstruction

techniques or the literature. In DCM, an experiment is considered as a perturbation of neuronal dynamics that is distributed through the brain network to generate source-specific responses. Experimental factors including stimulus property and task context change the parameters (i.e., effective connectivity) of the system. How can a forward or observational model best explain hidden neuronal activity? In the DCM for evoked responses, an inversion model considers data in both space and time. The parameters of a neuronal model include connectivity strength and time delays among sources. The spatial parameters include the location and orientation of each single source. With prior knowledge, one can also define more than one model with different sources and connections between sources, and then choose the best model or model family using Bayesian Model Selection (BMS) (see below).

Bayesian model inversion. For a given datum y and a specific DCM m , the inversion of m corresponds to approximating the posterior probability on the parameters θ given by Bayes' rule.

$$p(\theta|y,m) = \frac{p(y|\theta,m)p(\theta,m)}{p(y|m)}$$

where θ includes the parameters for forward, backward, and lateral connections and their modulation, which are to be estimated. The estimation procedure is fully described in Friston (2002). As Garrido et al. (2007) summarise, “This approximation uses variational Bayes that is formally identical to Expectation–Maximisation (EM), as described in Friston (2002). The EM can be formulated in analogy to statistical mechanics as a gradient descent on the free energy, F , of a system. The aim is to minimise the free energy with respect to a variational density $q(\theta)$. When the free energy is minimised $q(\theta) = p(\theta|y,m)$, the free energy $F = -\ln p(y|m)$ is the negative marginal log-likelihood or negative log-evidence. After convergence and minimisation of the free energy, the variational density is used as an

approximation to the desired conditional density and the log-evidence is used for model comparison.” (Garrido, Kilner, Kiebel, Stephan, & Friston, 2007, p. 573).

Bayesian model selection (BMS). BMS can be used to identify a model that best explains the observed responses. Different models are compared by their evidence (Penny et al., 2004). The evidence can be decomposed into two components: an accuracy term quantifying the data fit and a complexity term penalising models with a large number of parameters. Thus, a good model combines these two conflicting requirements to make it explain the data and keep it as simple as possible. The best model for a given dataset is the one with highest log-evidence $\ln p(y|m)$. For example, two models, m_1 and m_2 , the difference in their log-evidences $\ln p(y|m_1) - \ln p(y|m_2)$ or their Bayes factor (Penny, Stephan, Mechelli, & Friston, 2004) can be compared. If the difference is greater than 3, it can be concluded that the first model has the strongest evidence and thus is the best model. The BMS procedure is also applicable to select the best model family in which models with a specific feature are included (Penny et al., 2010).

Bayesian model averaging (BMA). The best model obtained by BMS can estimate effective connectivity between or within sources modulated by experimental manipulation (e.g., stimulus attributes or task demands). BMA is used to estimate the modulatory effects of an experimental manipulation, but not for a single model. BMA is used to average all family models with a specific model design, or to average models in a winning family after BMS is applied to family-level model selection (Penny et al., 2010). BMA will be used in the present thesis. A non-parametric proportion test is usually applied after BMA to test whether the statistical significance of an effective connection (e.g., Richardson, Seghier, Leff, Thomas, & Price, 2011; Xu, Wang, Chen, Fox, & Tan, 2015).

DCM for evoked responses is a powerful tool to elaborate millisecond-level temporal information propagation in brain networks, especially bottom-up and top-down directional

connections. It combines both spatial and temporal information, helping to enhance understanding of neural dynamics that underlie cognitive processes. Since the development of DCM for evoked responses (David et al., 2006), it has been widely used to examine effective connectivity in the M/EEG data (e.g., Garrido, Kilner, Kiebel, & Friston, 2007; Kiebel et al., 2009; Boly et al., 2011). Its widespread application in the cognitive neuroscience provides an innovative way to understand the neural mechanisms of human behaviours.

To conclude, DCM can serve as an ideal tool to examine backward (top-down) and forward (bottom-up) information flow in the brain when a task is performed. A large number of neuroimaging studies have demonstrated the value of DCM in unveiling inter-regional effective connectivity. DCM for evoked responses will be used in **Chapter 2** to examine top-down feedback from LIFG to LvOT at the early stages of visual word recognition. In the following section, I will introduce the time course of visual word recognition and explain the importance of examining early stage brain activity.

Time Course of Visual Word Recognition

Neuropsychological and fMRI data can help to unveil which brain regions respond to visual words and the functional roles of the connections between these regions in visual word recognition. However, the cognitive processes involved in visual word recognition unfold rapidly; important cognitive stages of word recognition can be completed in half a second (Serenio & Rayner, 2003; Carreiras et al., 2014; Hauk, 2016). Therefore, an investigation of early cognitive processing stages is largely beyond fMRI's scope because it has very poor temporal resolution. In contrast, non-invasive EEG and MEG techniques have excellent temporal resolution and thus can serve to measure early brain activity. EEG records electrical

activity of the brain by placing electrodes along the scalp, while MEG uses sensitive magnetometers to record magnetic fields produced by electrical currents occurring in the brain.

Due to this advantage, EEG and MEG can be used to elucidate the time course of visual word recognition (Sereno & Rayner, 2003; Carreiras et al., 2014; Hauk, 2016). Over the past two decades, researchers have found several event-related EEG or MEG components that are associated with visual word recognition. Basic visual processing of visual words occurs at around 100 ms after stimulus onset, which is followed by orthographic form analysis occurring at around 150 or 170 ms (e.g., Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Maurer, Brandeis, & McCandliss, 2005). Some investigations have shown that lexical-semantic access can take place within 200 ms after stimulus onset (e.g., Hauk et al., 2006; Hauk, Coutout, Holden, & Chen, 2012). For example, the differences in the neural responses between real words and pseudowords occur within 200 ms after stimulus onset (Hauk et al., 2006, 2012). A detailed description of these electrophysiological components of visual word recognition can be found in the reviews by Carreiras et al. (2014) and Grainger and Holcomb (2009).

With appropriate inverse algorithms, MEG sensor signals can also be localised to specific brain regions where these signals originate from. Studies with reading-related tasks have reported a consistent and coherent pattern of posterior-to-anterior activation in the brain. That is, the activation evoked by visual words starts from primary occipital cortex at ~100–130 ms, then moves to the ventral visual pathway in the left vOT at ~150–170 ms and to left-lateralised activity in the temporal and inferior frontal cortex from around 200 ms onwards (e.g., Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999; Pammer et al., 2004; Cornelissen et al., 2009; Vartiainen, Parviainen, & Salmelin, 2009; Wheat, Cornelissen, Frost, & Hansen, 2010; Klein et al., 2014). Recent years have also witnessed the interesting finding that there is a very early response to written words in the LIFG within 200 ms after stimulus

onset (e.g., Pammer et al., 2004; Cornelissen et al., 2009; Wheat et al., 2010; Klein et al., 2014). For example, using a lexical decision task, Pammer et al. (2004) found early activity in the LIFG for words in the 100-300 ms time-window.

How do distributed brain regions interact at the early stages of visual word recognition? This question still remains to be answered. A visual word cannot be recognised purely by the primary occipital cortex, but has to involve neural interactions between the occipital cortex and spoken language network. Thus, visual word recognition has to be completed on the basis of inter-regional interactions. A large number of fMRI studies have observed inter-regional interactions during reading and visual word recognition, especially between the LvOT and the fronto-temporal regions (Bitan et al., 2005, 2006; Heim et al., 2009). However, M/EEG studies are needed to explore in detail these interactions because visual word recognition is rapid.

One of these inter-regional interactions is the feedback from frontal spoken language regions to the LvOT. This feedback is likely to be crucial for the rapid retrieval of a word's sound and meaning because frontal spoken language regions are mainly responsible for processing phonological and lexical-semantic information. Although early activity of the LIFG during visual word recognition has been established (Pammer et al., 2004; Cornelissen et al., 2009), how the frontal language regions interact with the LvOT at the early stages of reading and visual word recognition is largely unclear. Using DCM for evoked responses, a recent MEG study found that the connection from LIFG to LvOT was stronger for visual words than for meaningless symbols within 200 ms after word onset (Woodhead et al., 2014). This finding indicates that the LIFG sends an early top-down feedback to the LvOT for processing word-specific properties. However, it is still unknown what the nature of this early top-down feedback is; to be specific, what unique responsibility this top-down feedback is taking at the early stages of visual word recognition. This question was not answered in Woodhead et al.

(2014). Therefore, it will be worth exploring the nature of the early top-down feedback from frontal language regions to the LvOT during visual word recognition.

In summary, studies with EEG and MEG techniques can add to the understanding of early neural responses during visual word recognition. Rapid neural activities elicited by visual words can be used to examine ongoing cognitive processing stages, which cannot be captured by BOLD signals due to poor temporal resolution. An outstanding question is how the frontal language regions interact with the LvOT at the early stages of visual word recognition, which will be examined in **Chapter 2**. In the following section, I will provide a review of task-dependent neural activity of visual word recognition, based on which I propose another outstanding question: how do task goals modulate the early top-down from LIFG to LvOT during visual word recognition?

Task-Dependent Neural Activity of Visual Word Recognition

Cognitive processes and neural correlates of visual word recognition are examined when a specific task, such as lexical decision or word naming, is performed. However, both cognitive and neural aspects of visual word recognition are highly modulated by task goals because different tasks recruit different cognitive and neural resources to serve specific goals.

Behavioural studies have revealed that the influence of linguistic variables, such as word frequency and imageability, on visual word recognition varies from task to task (e.g., Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004). For instance, word frequency effects contribute more to performance on a lexical decision task than a speeded naming task (Balota et al., 2004; Yap & Balota, 2009). Some fMRI studies have found that the neural basis of visual word processing is also modulated by task goals. This has been demonstrated in relation to the activation in local brain regions (e.g., Twomey, Duncan, Price, & Devlin, 2011; Yang, Wang,

Shu, & Zevin, 2012; Mano et al., 2013; McNorgan et al., 2015). For instance, in a meta-analysis of reading, McNorgan et al. (2015) found that pseudowords elicited more activation than words in the LvOT and LIFG, but words elicited more activation than pseudowords in the angular gyrus across lexical decision and naming tasks. More interestingly, the activation in the LvOT and the LIFG was higher in the naming task while the activation in the angular gyrus was higher in the lexical decision task. Some fMRI studies have further found that inter-regional connectivity during visual word processing is also modulated by task goals (e.g., Pugh et al., 2002; Bitan et al., 2005, 2006; Heim et al., 2009). For example, using a spelling task and a rhyming task, Bitan et al. (2005, 2006) found that the spelling task recruited stronger directional connections to the intra-parietal sulcus involving orthographical form analysis and the rhyming task recruited stronger directional connections to left temporal cortex involving phonological form analysis. Collectively, the existing findings provide strong evidence that regional activation and inter-regional connectivity on visual word recognition are not purely stimulus-driven, but highly task-guided instead: a task with greater emphasis on a specific linguistic property elicits more activation in some specific regions and connections than in others.

Do different task goals affect the early neural dynamics of visual word recognition? The answer appears to be yes. A large number of electrophysiological studies have examined whether task goals affect the time course of word recognition (e.g., Bentin et al., 1999; Strijkers, Yum, Grainger, & Holcomb, 2011; Strijkers, Bertrand, & Grainger, 2015; Chen, Davis, Pulvermuller, & Hauk, 2015; Mahé, Zesiger, & Laganaro, 2015; Wang & Maurer, 2017). Lexical decision, reading aloud, silent reading, and semantic categorisation are the most widely-used tasks. These studies have found differential influences of these different tasks on the time course of word recognition within the first 200 ms after stimulus onset. For example, Mahé et al. (2015) found that the electrophysiological responses evoked by words diverged

between reading aloud and lexical decision tasks from about 140 ms. Strijkers et al. (2011) found that word-evoked responses between a reading aloud task and a semantic categorisation task diverged at around 170 ms. Strijkers et al. (2015) found that word frequency effects reflecting lexical-semantic access occurred at 120 ms onward during a semantic categorisation task, but occurred at 220 ms during a colour categorisation task.

These findings demonstrate that tasks with different cognitive goals modulate the time-course of visual word recognition even at early stages (e.g., the first 200 ms after stimulus onset). This leads to a reasonable expectation that the top-down feedback from LIFG to LvOT within 200 ms after stimulus onset is likely to be also task-dependent, although this question remains to be answered. Building on Chapter 2, **Chapter 3** will address this issue by using a task different from that in Chapter 2 and comparing findings between the two tasks. The existing electrophysiological studies examining task modulation of the time course of visual word recognition are limited in that the focus of these studies was mainly on single time-window or single linguistic property (e.g., Bentin et al., 1999; Strijkers et al., 2015; Wang & Maurer, 2017). Using these studies as a starting point, Chapter 4 will examine task modulation of the time course of visual word recognition for different linguistic properties in multiple time-windows.

Outline of This Thesis

The detailed introduction provided of the neural basis of visual word recognition reveals that the journey from visual features to meaning for visual word recognition is supported by local regions distributed in the brain and the neural interactions between them. For skilled readers, cognitive processes of visual word recognition are fast-acting; thus, it is important to examine early neural dynamics. Electrophysiological studies have demonstrated

that within 200 ms after stimulus onset, linguistic properties of visual words can be rapidly accessed and this early processing is highly modulated by task goals. More interestingly, within 200 ms after stimulus onset, the LIFG is reliably activated and this frontal region sends top-down feedback to the LvOT. However, it is largely unknown what the nature of this early top-down feedback from LIFG to LvOT is (outstanding question 1), and how task goals modulate this early feedback (outstanding question 2). In addition, in terms of the automaticity of visual word recognition, an outstanding question is how task goals influence the time course of visual word recognition, specifically in a non-linguistic task with no requirement for any linguistic processing (outstanding question 3). Using MEG with excellent temporal resolution, the current thesis sought to investigate examine these outstanding questions about the early neural dynamics of visual word recognition in three different studies. The first study explored in detail the nature of early top-down feedback from LIFG to LvOT. The second study examined the task dependence of this early top-down feedback. The investigations of neural connectivity between local regions offered us an opportunity to look at how event-related fields evoked by visual words unfold over time. With the focus on event-related fields, the third study aimed to examine task modulation of the time course of visual word recognition. We addressed these research questions in three individual chapters of the thesis, presented in journal article format.

Building on Woodhead et al. (2014) who found an early top-down feedback from LIFG to LvOT for visual words, **Chapter 2** presents the first empirical study that aimed to examine the nature of this early top-down feedback during visual word recognition. Four types of stimuli - real words, pseudowords, consonant strings and false fonts - were used to construct three comparisons, real words vs. pseudowords (lexicality effects), pseudowords vs. consonant strings (phonological effects), and consonant strings vs. false fonts (letter effects). In this MEG study, fifteen young adults performed a semantic categorisation task, and the DCM for evoked responses was used in these three different comparisons to model inter-regional causal

connections. It was expected that lexicality and phonological effects would take place in the first 200 ms because the LIFG is a region mainly used for high-level linguistic processing.

Chapter 3 presents a second empirical study that aimed to examine how task goals modulate the early top-down feedback from LIFG to LvOT. In contrast to the semantic categorisation task, a visual discrimination task was employed, in which participants were asked to detect hash strings. Similar to the previous study, the DCM for evoked responses was used in these three different comparisons to model inter-regional causal connections. It was expected that this early top-down feedback would be altered in the context of the visual discrimination task.

Chapter 4 presents a study that combined the two MEG experiments in Chapters 2 and 3 but adopted a different method from DCM to examine task-modulation of the time course of visual word recognition. The focus was on the three comparisons defined above. In addition, a behavioural version of each MEG experiment was also carried out to allow comparison between neural and behavioural data. Considering the nature of each task, it was expected that the neural and behavioural differences between conditions would be amplified in the semantic categorisation task, but would be weakened, delayed or disappear in the visual discrimination task.

Chapter 5 presents a general discussion. It provides a summary of the main findings and discusses the findings from each empirical study by linking them to previous findings and integrating them within a framework of early neural dynamics of visual word recognition. Potential implications from this thesis to related topics and several limitations are considered.

Chapter 2

Early Top-down Feedback from Frontal to Ventral Occipito-Temporal Cortex during Visual Word Recognition

Yu Li, Sachiko Kinoshita, Paul Sowman, & Anne Castles

ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney

Note: This chapter was presented at at the 24th Annual Meeting of the Cognitive Neuroscience Society, San Francisco, CA, United States.

Abstract

Visual word recognition involves linking written symbols to their spoken forms and lexical-semantic representations. In the brain, this process is built on the inter-regional connections between the left ventral occipito-temporal cortex (LvOT) for orthographic form analysis and the left inferior frontal gyrus (LIFG) for lexical-semantic and phonological analysis. Building on a previous magnetoencephalography (MEG) study showing that the LIFG-to-LvOT connection was stronger for words than for meaningless false fonts within 200 ms of visual word recognition (Woodhead et al., 2014), the current study aimed to examine the nature of this early top-down feedback by applying dynamic causal modeling. A group of 15 adult participants performed a semantic categorisation task. Four comparisons - real words vs. false fonts, real words vs. pseudowords, pseudowords vs. consonant strings, and consonant strings vs. false fonts - were used in the 1-100, 1-150 and 1-200 ms time-windows to examine the time course of the influences of lexical-semantic, phonological and letter information on the early top-down feedback. The results of Bayesian Model Averaging showed that (1) the LIFG-to-LvOT connection was stronger for real words than for false fonts in the 1-150 ms time-window, replicating the previous observation in Woodhead et al. (2014); (2) it was stronger for real words than for pseudowords in both the 1-150 ms and 1-200 ms time-windows, indicating an early lexicality effect; (3) it was stronger for pseudowords than for consonant strings in the 1-200 ms time-window, indicating a phonological effect which occurred later than the lexicality effect; and (4) it was also stronger for false fonts than for consonant strings while the right homologue (the RIFG-to-RvOT connection) was stronger for consonant strings than for false fonts in both the 1-150 ms and 1-200 ms time-windows, indicating letter effects. These results suggest that compared with phonological information, lexical-semantic and letter information in words may exert earlier influences on the top-down feedback from LIFG to LvOT.

Introduction

The writing system - the visual form of oral language to be used in recording facts and transferring thoughts or ideas - was invented only about 5400 years ago (Robinson, 2009). It is unlikely that this is sufficient time for the brain to evolve a specific area for written language processing. Therefore, it is generally assumed that reading ability is not innate, and involves recruiting existing structures of the human brain that have evolved for other purposes (Dehaene & Cohen, 2007). There is mounting neuroimaging evidence that reading involves several distributed areas in the brain including the left inferior frontal gyrus (Broca's area), the left posterior superior temporal cortex (Wernicke's area) and the visual cortex. It has been also consistently demonstrated that there is a region along the ventral pathway of visual processing that is always activated more by words than by meaningless symbols presented visually (McCandliss, Cohen, & Dehaene, 2003; Dehaene & Cohen, 2011). This area is located in the left ventral occipito-temporal cortex (vOT) and widely thought of as a crucial area for processing visual word forms (McCandliss et al., 2003; Dehaene & Cohen, 2011). Furthermore, there is also evidence that the left vOT receives "top-down" influences from the inferior frontal language area (Woodhead et al., 2014). The aim of the present study was to examine in detail the nature of these top-down influences from the frontal language area to the left vOT.

As a process of recovering language from visual symbols, reading involves making connections between visual forms of written words and their spoken forms and meanings. In the brain, these connections are established via pathways between the vOT and areas of the spoken language network including the articulatory area the inferior frontal gyrus (IFG). Indeed, increasing magnetic resonance imaging (MRI) studies documents white matter pathways between vOT and spoken language areas (Ben-Shachar, Dougherty, & Wandell, 2007; Yeatman, Rauschecker, & Wandell, 2013). Evidence of functional connectivity between

these areas has also come from reports of spontaneous neural synchronisations between the vOT and language areas (Koyama et al., 2011; Li et al., 2017). Concerning the nature of the functional role of the left vOT in visual word recognition, there are two main accounts - the Local Combination Detector model (LCD) proposed by Dehaene, Cohen, Sigman, and Vinkier (2005) and the Interactive Account proposed by Price and Devlin (2011), which mainly differ on whether higher-level language areas provide top-down modulation of the activity in the left vOT. The LCD model argues that the left vOT is a pre-lexical hub, specific for written words, that computes and stores strictly visual and abstract pre-lexical orthographic representations in a primarily feed-forward manner; this area was thus labelled as the visual word form area (VWFA) (Cohen et al., 2000; Dehaene et al., 2005; Dehaene & Cohen, 2011). In contrast, the Interactive Account holds that the vOT is not specific to written words, but also responds to other kinds of stimuli such as pictures. More importantly, the activation in the left vOT is proposed to be modulated by higher-level regions which are engaged in linguistic processing of written words; the left vOT continuously and automatically interacts with other regions during reading, acting as an interface linking bottom-up visual form information critical for orthographic processing with top-down high-level linguistic properties of written words (Price & Devlin, 2003, 2011).

Findings from a large number of fMRI studies support the Interactive Account. Specifically, activity in the left vOT cortex has been shown to be sensitive to manipulations of task context (e.g., lexical decision task, reading aloud and passive viewing that emphasise different cognitive processes of reading) or high-level stimulus properties (e.g., letter strings with and without meaning) that are best explained by top-down feedback (Starrfelt & Gerlach 2007; Cai, Lavidor, Brysbaert, Paulignan, & Nazir, 2008; Van der Haegen, Cai, & Brysbaert, 2012; Kherif, Josse, & Price. 2011; Twomey et al. 2011; Mano et al., 2013). For example, the lateralisation of vOT activity during word reading highly depends on the lateralisation of

inferior frontal gyrus activity during word generation (Cai et al., 2008; Van der Haegen et al., 2010), indicating that this area for visual word form processing is influenced strongly by the high-order spoken language area. Mano et al. (2013) showed that subregions within the left posterior occipito-temporal sulcus are more activated by real words and pseudowords than by consonant strings when participants are asked to name visual stimuli. Twomey et al. (2011) argued that both stimulus- and task-driven modulations of activation in the left vOT can only be explained by top-down processing of non-visual aspects of tasks and stimuli, which are consistent with the hypothesis that the left vOT acts as an interface linking visual forms with linguistic processing in both bottom-up and top-down manners.

A limitation of the studies described above is that they relied on the blood-oxygen-level dependent (BOLD) signals recorded by fMRI techniques, which are limited by their relatively poor temporal resolution (Hall, Robson, Morris, & Brookes, 2014). The BOLD signal that these studies used to infer underlying neuronal activity in local areas is slow to emerge. The time course of brain activity is ultra-rapid and the cognitive processing stages (e.g., lexical access) reflected in the neural information flow are very fast (Hauk, 2016), which is beyond the scope of fMRI techniques. Magneto- and electro-encephalography techniques (M/EEG) with high temporal resolution have been increasingly used to investigate the time course of neural activation during cognitive tasks. M/EEG studies have found that one of the earliest markers of visual word recognition is a left-lateralised response termed the M/N170 that occurs around 170 ms after visual word onset and mainly reflects orthographic form processing (Maurer, Brandeis, & McCandliss, 2005; Hsu, Lee & Marantz, 2011; Dunabeitia, Dimitropoulou, Grainger, Hernandez, & Carreiras, 2012; Carreiras, Armstrong, Perea, & Frost, 2014), and that the influence of high-level linguistic information (e.g., lexical and semantic representations) occurs before 200 ms (e.g., Assadollahi & Pulvermüller, 2003; Sereno & Rayner, 2003; Carreiras, Vergara, & Barber, 2005; Dambacher, Kliegl, Hofmann, & Jacobs,

2006; Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson, 2006; Hauk, Coutout, Holden, & Chen, 2012). Due to their relatively high temporal resolution, electrophysiological measures are an ideal tool to study the time course of visual word recognition and the early modulations of lower-level visual feature detection, orthographic form analysis, and higher-level lexical and semantic information on word recognition.

M/EEG studies with reading-related tasks have reported a consistent and coherent pattern of posterior-to-anterior activation in the brain. That is, from primary occipital cortex (OCC: ~100–130 ms), then to the ventral visual stream with a peak in the left vOT (~150–170 ms) and left-lateralised activity in the temporal and inferior frontal cortex from around 200 ms onwards (Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999; Marinkovic et al., 2003; Pammer et al., 2004; Pykkänen & McElree, 2007; Cornelissen et al., 2009; Vartiainen, Parviainen, & Salmelin, 2009; Wheat, Cornelissen, Frost, & Hansen, 2010; Chen, Davis, Pulvermuller, & Hauk, 2013; Klein et al., 2014). The most interesting finding is the very early response to written words in the left inferior frontal gyrus (LIFG) within 300 ms post-stimulus (e.g., Pammer et al., 2004; Cornelissen et al., 2009; Wheat et al., 2010; Klein et al., 2014). For example, Cornelissen et al. (2009) found a left-lateralised inferior frontal gyrus response to words at 100-250 ms that was stronger than the response to consonant strings during a passive viewing task in which participants looked at stimuli one by one but had no explicit task. The result is in line with previous studies using different tasks (e.g., Pammer et al., 2004 in which participants were asked to decide whether a visual word presented was familiar).

In a recent MEG study, Woodhead et al (2014) used dynamic causal modeling (DCM), an analysis method that can identify the influences of one brain area on another, to specifically test the hypothesis that the activation in the LIFG might provide early feedback to the LvOT during visual word processing. In Woodhead et al (2014) study, real words and strings of false fonts (meaningless symbols) were used as critical stimuli, and personal names were used as

catch trials; participants were asked to read words and to view false fonts silently, and to respond by pressing a button when a person name was presented. The results showed that during the first 200 ms after stimulus onset, the connection from the left IFG to the left vOT was stronger for real words than for false fonts. This finding is broadly consistent with the assumption of the Interactive Account model of visual word recognition that the LvOT is an interface linking bottom-up visual inputs from occipital cortex and top-down linguistic influences from language areas (Price & Devin, 2011; Carreiras et al., 2014). Woodhead et al. argued that the top-down feedback they observed may reflect that the LIFG provides fast phonological cues to constrain visual feature processing from LIFG to LvOT, and this argument is consistent with the findings from two other MEG studies (Wheat et al., 2010; Klein et al., 2014). Specifically, Wheat et al. (2010) using a pseudohomophone masked priming paradigm found that the induced activation in the LIFG was significantly stronger when targets and primes had the same pronunciations than when they were different, supporting a functional role of the site in pre-lexical access to phonological information. Klein et al. (2014) found that the early activation in the LIFG was higher when the task involves articulation, suggesting that the early activity in that region may reflect fast access to phonological and articulatory codes.

However, direct evidence that the top-down feedback from LIFG to LvOT reflects rapid phonological activation is still lacking. The two studies described above reported evidence for a phonological locus of the activation in the LIFG but do not provide any evidence that it has a modulating effect on the LvOT. And the findings of the only study that did report evidence for a top-down feedback from LIFG to LvOT – Woodhead et al. (2014) – do not establish that the early influences from LIFG to LvOT are phonological in nature. There are at least three alternative accounts of the early top-down feedback reported by Woodhead et al. (2014) based on their comparison of real words versus false fonts: these could be lexical-semantic, phonological or even low-level letter effects. Real words have a meaning, are

pronounceable, and are made up of a sequence of familiar letters, whereas false fonts are strings of symbols without meaning for which a pronunciation cannot be generated. The existing studies has also implicated the LIFG in multiple aspects of reading and visual word recognition including lexical-semantic processing, phonological retrieval and manipulation, and vocal response (e.g., Wheat et al., 2010; Price, 2012; Taylor, Rastle, & Davis, 2013; Klein et al., 2014), thus each of these linguistic aspects represents a potential source of the top-down modulation.

The aim of our study was to examine in detail the nature of the early feedback from LIFG to LvOT during visual word recognition by applying the DCM technique in MEG data. To achieve this, we modified the design of Woodhead et al.'s (2014) study to include stimuli that would allow us to distinguish between a high-level lexicality effect, a phonological effect, and a low-level letter effect. Specifically, four types of stimuli were chosen: (1) real words (meaningful and pronounceable, consisting of letters), (2) pseudowords (meaningless, pronounceable and orthographically-legal, consisting of letters), (3) consonant strings (meaningless and unpronounceable, consisting of letters), and (4) false fonts (meaningless and unpronounceable, consisting of non-letters). The comparison between real words and pseudowords allows us to examine whether lexical and/or semantic linguistic information influences early top-down feedback. This is because skilled readers have lexical entries for real words and have memory traces of these words and corresponding meanings, but have no such corresponding lexical knowledge about pseudowords. Therefore, by comparing early top-down feedback effects for words with that for pseudowords which are orthographically legal and pronounceable, any specific influences of lexical factors can be identified. The comparison between pseudowords and consonant strings allows us to examine whether the phonological aspects of a letter string affect the top-down feedback. The phonology of pseudowords can be constructed through the application of grapheme-to-phoneme conversion, but this is not

possible for the unpronounceable consonant strings, thus any differences between these two conditions can be attributed to non-lexical phonology. Finally, the comparison between consonant strings and false fonts allows us to identify any influences that are attributable to letter recognition processes, as these two conditions differ in that consonant strings are formed by letters while false fonts are formed by non-letter visual unfamiliar symbols. These three different comparisons were used to partition different effects with distinct linguistic properties (see a similar design in Coch, 2015; Coch & Meade 2016), with the aim of specifying the nature of the early top-down feedback of the LIFG to the LvOT. Following an approach in previous studies (Garrido, Kilner, Kiebel, & Friston, 2007; Woodhead et al., 2014; Poch et al., 2015), the current study used three time-windows - 1-100, 1-150, and 1-200 ms - in the DCM analysis to explore the time frame of this early top-down feedback.

Increasing electrophysiological studies have found that real words can be clearly discriminated from pseudowords at around 170 ms (e.g., Sereno, Rayner & Posner, 1998; Maurer et al., 2005; Hauk et al., 2006, 2012; Mahé, Zesiger, & Laganaro, 2015; Araújo, Faísca, Bramão, Reis, & Petersson, 2015; Coch & Meade, 2016), indicating early lexicality effects. Lexical variables, such as word frequency, also were found to affect the first 200 ms brain activity (e.g., Assadollahi & Pulvermuller, 2001, 2003; Sereno et al., 1998; Hauk & Pulvermuller, 2004; Mahé et al., 2012; Strijkers, Bertrand, & Grainger, 2015). Assadollahi & Pulvermuller (2001, 2003) found that a word frequency effect was observed at around 150 ms. But there is also evidence that lexicality effects take place at around 400 ms (M/N400 effect, e.g., Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999). Interestingly, some studies have found that pseudowords and consonant strings can be differentiated from each other in the first 200 ms (e.g., Martin et al., 2006; Coch & Mitra, 2010; Araújo et al., 2015) and the rapid differentiation of consonant strings from false fonts in the first 200 ms (Bentin et al., 1999; Mahé et al., 2012; Araújo et al., 2015; Coch & Meade, 2016), indicating early

phonological and letter effects. However, the electrophysiological evidence above cannot provide direct inferences about early effective connectivity between the frontal cortex and ventral occipito-temporal cortex. Therefore, no specific expectations were made in this study. Instead, it was generally expected that both lexicality and phonological effects would be observed in the top-down feedback in the first 200 ms after stimulus onset; but letter effects would not be observed in this top-down feedback because the LIFG is mainly responsible for high-level lexico-semantic and phonological processing (e.g., Taylor et al., 2013; Carreiras et al., 2014).

Methods

Participants

Fifteen native English speakers (7 females, 9 males; age range: 19-28 years; mean age: 22.38 years) were recruited from Macquarie university campus to participate in the study. All participants reported that they had normal hearing, normal or corrected-to-normal vision and had no history of neurological disorders. The Edinburgh Handedness Inventory (Oldfield, 1971) was used to assess their handedness: fourteen were right-handed and one left-handed. The Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) subscales of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner & Rashotte, 1999) were used to confirm that all participants had reading competence within the normal range (mean of overall standard score: 104 ± 13.38). All participants had normal performance IQ (mean of overall standard score: 118 ± 8.42 ; range: 103-132) as measured by the Matrices subscale of the Kaufman Brief Intelligence Test, (Second Edition; KBIT-2; Kaufman & Kaufman, 2004). The study was approved by the Human Research Ethics Committee (Medical Sciences) at

Macquarie University. Prior to the beginning of the experiment, a signed consent form was obtained from each participant.

Overall Procedure

Participants received the 80-minute MEG experiment first and then the 30-minute behavioural testing including the reading ability test (TOWRE) and the intelligence test (KBIT) either on the same day or the following day.

MEG Experiment Materials

Visual stimuli belonged to four categories: real word (RW), pseudoword (PW), consonant strings (CS) and false fonts (FF) (see examples in Figure 1). Within each category there were 120 exemplars divided equally into four stimulus lengths of three, four, five and six characters. All stimuli were presented in Calibri in lower case and size 50. RWs were chosen from MCWord (an Orthographic Word form Database; for details, see <http://www.neuro.mcw.edu/mcword/>). The MCWord is a database based on the CELEX efw.cd file that contains approximately 16,600,000 examples. PWs were first generated based on real words through the WUGGY software (Keuleers & Brysbaert, 2010. For details, see <http://crr.ugent.be/programs-data/wuggy>) and then were carefully chosen to match RWs in bigram frequency, trigram frequency, and Coltheart's N based on the MCWord database. Two-sample t-tests revealed no significant differences between RW and PW for each variable ($p > 0.10$); see Table 1 for details. CSs were also initially generated by the WUGGY software, with bigram/trigram frequency and Coltheart's N set at zero to keep CSs as pure letter strings without any higher-level orthographic information. As in the previous study on which this study builds (Woodhead et al., 2014), FFs were direct translations of the real words using the "Carian" font (Jane Warren, personal communication). "Carian" fonts are characters adapted

from the alphabet of an obsolete Anatolian language (Melchert, 2004). To reduce possible influences of similarity to English letters, some characters whose shape was visually similar to English letters were replaced by others. Therefore, the FFs had no meanings and could not be pronounced.

Table 1.

Psycholinguistic properties including word frequency, Coltheart's N, Bigram and trigram frequencies for real word and pseudowords.

	Real words			Pseudowords		
	mean	SD	range	mean	SD	range
Word frequency (per million)	77.25	70.18	8.80-268.07	N/A	N/A	N/A
Coltheart's N (number of neighbors)	6.55	6.19	0-23.00	6.38	5.85	0-21
Bigram frequency (per million)	882.87	624.21	99.23-3015.16	854.91	603.07	73.92-2971.28
Trigram frequency (per million)	199.10	193.34	11.01-1153.94	162.98	180.82	0-952.64

Critical conditions				Targets
real words	pseudowords	consonant strings	false fonts	animal words
guy	sog	qvr	ᵀᵀᵀᵀ	ant
camp	gasy	lxhp	ᵀᵀᵀᵀ	frog
blood	aloze	hhfzt	ᵀᵀᵀᵀ	hippo
honour	myntem	nsqrjk	ᵀᵀᵀᵀ	cicada

Figure 1. Left, examples of critical stimuli, real words, pseudowords, consonant strings, and false fonts; right, examples of target stimuli (animal words).

MEG Experiment Procedure

The entire MEG experiment procedure included preparation, a practice run, the actual experiment, and rests between blocks. The actual experiment consisted of six blocks. Each

block lasted about 6 minutes. Each block contained 96 trials - 20 RWs, 20 PWs, 20 CSs, 20 FF - and 16 animal words (AW, e.g. *cat*, *dog* and *donkey*) which functioned as “catch trials”. Thus, the experiment included 480 trials (120 trials for each condition) and 96 animal word trials. The order of stimuli was pseudo-randomised and presented on the screen in the magnetically shielded room to the participants using E-Prime 2.0 (<https://www.pstnet.com/eprime.cfm>). The distance between the screen and participants’ eyes was 60 cm. On each trial, a black screen was initially presented with a white cross for visual fixation; then a stimulus was presented for 500 ms, followed by a white cross which was displayed for between 2800 ms and 3400 ms (see Figure 2). Participants were asked to attend to each trial carefully and respond with a button press when an animal name was presented on the screen. Catch trials were removed from the data analysis. Before the experiment, a practice run was conducted to familiarise participants with the task.

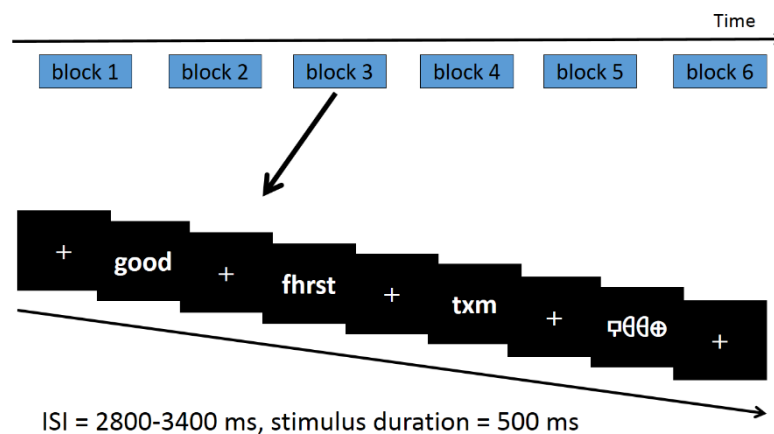


Figure 2. Experimental procedure for the study. Six blocks were included, with each having 90 trials including 80 critical trials and 16 animal trials.

MEG Data Acquisition and Pre-processing

MEG data were recorded using the KIT-Macquarie MEG160 (Model PQ1160R-N2, KIT, Kanazawa, Japan) located at Macquarie University, with participants lying in a

magnetically shielded room. Data were recorded using 160 coaxial first-order gradiometers with a 50-mm baseline (Kado et al. 1999; Uehara et al. 2003). Prior to MEG data acquisition, the locations of three head landmarks (nasion and bilateral preauricular points), five marker coil positions, and the subject's head shape were recorded by 3D digitisation (Polhemus Fastrack, Colchester, VT). Participants' head position and corresponding sensor positions were measured by energising the five marker coils briefly. The head motion was less than 6 mm for each participant in each block during the entire recording. The online sample rate was 1000 Hz.

MEG data analysis was conducted using SPM12 software (Litvak et al., 2011) implemented in Matlab 2014b (The MathWorks, Inc.). The original data sampled at 1000 Hz were high-pass filtered at 1 Hz and then low-pass filtered at 30 Hz. The filtered data were epoched from 150 ms before stimulus presentation to 600 ms after stimulus presentation. The pre-stimulus time window (-100 - 0 ms) was used for baseline correction. To reject artefacts, the Fieldtrip visual artefact rejection toolbox, which expresses every time point as a deviation from the mean over all time and channels, was used to remove extreme trials in four conditions for each participant based on the variance within each channel (http://www.fieldtriptoolbox.org/reference/ft_rejectvisual). No more than 6% of trials in any condition were excluded by the artefact rejection procedure. Subsequently, robust averaging was applied to the epoched data across trials within each condition (RW, PW, CS and FF) (Litvak et al., 2011; Wager, Keller, Lacey, & Jonides, 2005).

A head model for constructing source activity in the DCM analysis for each participant was built based on several settings below. First, a normal-resolution cortical mesh generated from the MNI template (ICBM512) was created. Then, three MEG fiducial labels (nasion, left pre-auricular point (LPA) and right pre-auricular point (RPA)) and headshape points were used to coregister the MEG data. Finally, a single shell was adopted to define the forward model in the following DCM analysis.

Dynamic Causal Modeling (DCM)

DCM is a widely-used and powerful tool for examining the causal relationships between brain areas (effective connectivity; the influence a region exerts on another region). DCM was first introduced for fMRI data (Friston, Harrison, & Penny, 2003) and then for M/EEG data (David et al., 2006). The central idea of DCM is that neuronal activity flows through brain networks in an input-states-output dynamical system. Inputs such as stimulus attributes and task demands evoke neural responses, hidden states include neuronal activities that form the outputs, and outputs are observable responses. A neuronal model with inputs and outputs is augmented with an observation model describing the mapping from hidden neuronal activity to observed responses. DCM starts with a neuronal model with interactions among cortical regions. The neuronal model's parameters, namely effective connectivity within or between regions, are then estimated from observed responses and how these parameters are influenced by experimental factors. Multiple models varying in nodes and connections can be designed to map measured data. Using a Bayesian model selection (BMS) procedure, one can get the best model or model family with a specific feature that most closely matches the measured data. Bayesian model averaging (BMA) can be used to average inter-regional effective connectivity across more than one model.

Here we used DCM for evoked responses. The details and basic principles of this have been extensively described elsewhere (David et al., 2006; Kiebel, David, & Friston, 2006; Kiebel, Garrido, Moran, Chen, & Friston, 2009). In DCM for evoked responses, a neural mass model (Jansen & Rit, 1995) is used to describe the neuronal dynamics of each single source. In the model, each source comprises three neuronal subpopulations. Each source is described by the average post-membrane potentials and mean firing rates of the three neuronal subpopulations. Mean firing rates from other sources arrive at a source via directed forward (or

bottom-up), backward (or top-down), and lateral connections. All these extrinsic cortico-cortical connections are excitatory and are mediated through the axons of pyramidal cells.

Basic steps of DCM. A standard pipeline of DCM for evoked responses analysis includes the following parameters that need to be specified: the definition of the source locations; the definition of the connections between these sources; the source(s) of input(s); the connections to be modulated by variations in stimulus types and/or tasks. A DCM model can be established through changing these parameters iteratively until the predicted neural activity produced by the model best matches the observed data. This also means that one can use this approach to estimate a large body of models by changing the number of sources, the connections between sources and sensory inputs, to test a specific hypothesis. Bayesian statistics are performed to investigate which model provides the best explanation of the data (Bayesian model selection, BMS; Penny et al., 2004). If there is no model that is apparently the best one, then a weighted average model can be established via the approach of Bayesian Model Averaging (BMA; Penny et al., 2010). In our case, here we used a BMA approach to estimate connection strengths.

Nodes of DCM. A meta-analysis based on eight MEG/EEG studies of visual word recognition (Pammer et al., 2004; Cornelissen et al., 2009; Wheat et al., 2010; Thesen et al., 2012; Yvert, Perrone-Bertolotti, Baciú, & David, 2012; Klein et al., 2014; Quinn, 2014; Woodhead et al., 2014. See supplementary table 1 for details), which reported coordinates of regions activated in the first 300 ms after stimulus onset, was performed using Ginger ALE software (<http://www.brainmap.org/ale/>). Multiple comparison correction was performed using False Discovery Rate (FDR; Laird et al., 2005; Genovese, Lazar, & Nichols, 2002). Three left-hemisphere regions were then obtained: left middle occipital cortex (OCC), left ventral occipito-temporal cortex (vOT) and left inferior frontal gyrus (IFG) (see Figure 3). The vOT coordinates were very close to those reported by previous studies (e.g., Cohen, Jobert, Le

Bihan, & Dehaene, 2004; Dehaene & Cohen, 2011). The right-hemisphere homologues of these areas were also obtained based on these coordinates. Thus, a six-node network for the DCM analysis was constructed which included bilateral OCC (MNI, ± 26 -94 6), bilateral vOT (MNI, ± 46 -56 -18) and bilateral IFG (MNI, ± 58 8 22). A study with similar experimental settings has shown that this six-node network provides the best fit to the MEG data (Woodhead et al., 2014), thus its nodes were considered as the sources of the DCM spatial model in the following BMA analysis.

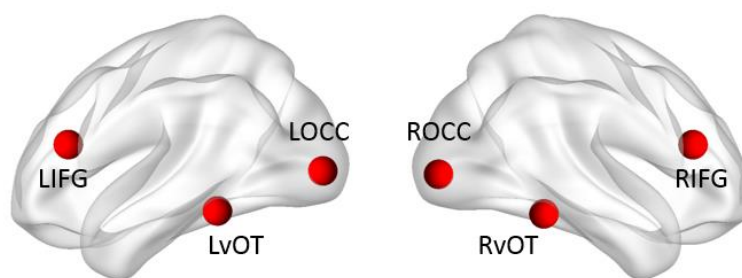


Figure 3. Six nodes (sphere in red) in the left and right hemisphere generated by an ALE meta-analysis based on eight MEG studies of reading.

DCM models. In the current model family, the total number of possible connections in a model with 6 nodes is 30, and the total number of possible combinations of those connections (the total model space) is 2^{30} . To reduce computational load, some constraints had to be enacted to limit the size of the model space. Following the previous most comparable DCM study (Woodhead et al., 2014), the model space was constrained using four rules. First, horizontal lateral connections within a level of the cortical hierarchy (e.g., from left OCC to right OCC) were allowed while diagonal lateral connections (i.e., lateral connections between levels of the cortical hierarchy, e.g., from left OCC to right vOT) were not allowed. Second, forward or backward connections (e.g., from left OCC to left vOT) were mirrored in the opposite hemisphere (right OCC to right vOT). Third, the connections between bilateral OCC and bilateral vOT were consistently kept on across all models. Fourth, lateral connections (e.g., left OCC to right OCC) had corresponding reciprocal connections (e.g., right OCC to left OCC).

With these rules, the number of independent connections was reduced to 5 in the current model family, creating a total model space of 32 models for each comparison for each participant; each modeling a different combination of connections between sources mediating trial-specific effects.

To better examine the changes in connection strength over time, the DCM analysis was conducted in three different time-windows, 1-100 ms, 1-150 ms and 1–200 ms (see Garrido et al. 2007, Woodhead et al., 2014 and Poch et al., 2015 for a similar approach). Three peaks at around 100 ms, 150 ms and 200 ms are also very clear in the early time-windows. These three time-windows were chosen based on previous studies in which main reading areas in the left hemisphere were clearly activated in the first 200 ms during visual word recognition (e.g., Pammer et al., 2004; Cornelissen et al., 2009; Wheat et al., 2010; Woodhead et al., 2014).

Bayesian Model Averaging (BMA). All 32 models in the model family were estimated in three time-windows for each participant. The BMA analysis was performed to estimate the strength of inter-regional causal connections in the six-node network. The 32 models were averaged by group-level BMA with random effects (Penny et al., 2010) to investigate the average strength of the stimulus-modulated connections, that is, the gains for the three comparisons: RW vs. PW, PW vs. CS, and CS vs. FF. The comparison RW vs. FF was also evaluated to provide an overall estimation and to replicate the finding reported by Woodhead et al. (2014) that RW, compared with FF, evoked a stronger connection from LIFG to LvOT in the first 200 ms after stimulus onset. For each connection, an average gain value equal to 0 indicates that no stimulus-type effect occurs in the connection; an average gain value greater than 0 indicates that the connection strength is stronger for the former stimulus type; and vice versa, an average gain value smaller than 0 indicates that the connection strength is stronger for the latter stimulus type.

A non-parametric proportion test was adopted to test whether the gain values from the BMA analysis were significant or not. For each connection, the distribution of the gain was reconstructed by generating 100,000 samples from a Gaussian distribution based on the posterior mean and standard deviation calculated in the BMA step. The rule was that if more than 90 percent of the samples were greater or smaller than zero, the connection was considered to be significantly stronger for the former condition or the latter condition, respectively; otherwise, the connection strength was considered to be equal between two conditions in the connection. For instance, in the comparison of RW vs. PW, if the gain of a connection is 0.30 and is significantly greater than 0, then the connection strength is stronger for RW. A similar approach has been used before (Seghier, Josse, Leff, & Price, 2011; Richardson, Seghier, Leff, Thomas, & Price, 2011; Woodhead et al., 2014; Xu, Wang, Chen, Fox, & Tan, 2015).

Results

The results of task performance and event-related field at the sensor space are reported below, and then followed by the results of the DCM analysis for each comparison.

Task Performance

The average reaction time for catch trials with correct response, i.e. animal words, was 659 ms ($SD = 89$ ms, range 554-882 ms). The average proportion accuracy for catch trials was 0.94 ($SD = 0.39$, range 0.85-0.99).

Event-related Fields (ERFs)

Root mean square (RMS) value was used to depict the time course of event-related fields globally across the sensor space. The RMS value is the square root of the arithmetic mean of the squares of the values in all MEG sensors at a given time-point and reflects how

strong an ERF is. The average RMS values across four conditions were calculated in the first 500 ms after stimulus presentation. As can be seen in Figure 4, event-related peaks at around 100 ms, 150 ms and 200 ms were clearly shown as a function of time-course. The observation is in line with previous studies of visual word recognition (e.g., Hauk et al., 2006; Chen et al., 2013).

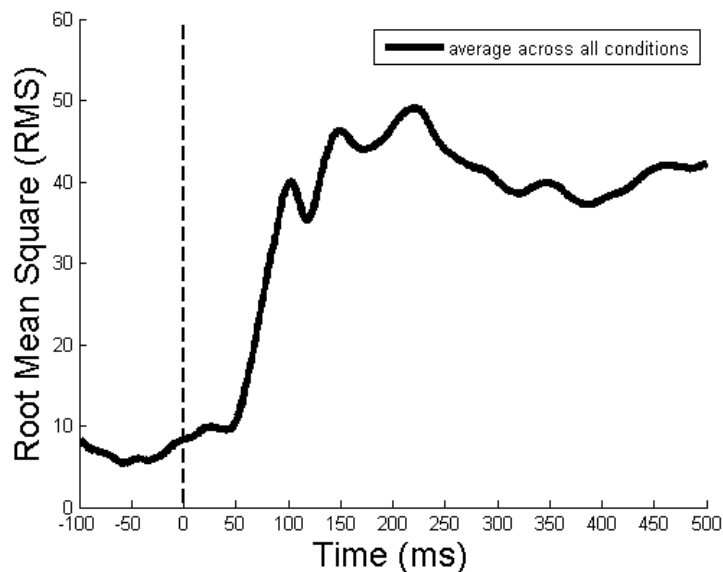


Figure 4. The average root mean square (RMS) plot across all conditions and all subjects as a function of time course during the first 500 ms after stimulus onset. Four event-related field peaks can be seen at 100 ms, 150 ms and 200 ms, i.e., M100, M150 and M200 components.

Bayesian Model Averaging (BMA)

All 32 models were included in the current DCM-BMA analysis. The results of the BMA analysis assessing the stimulus-type effects on the causal relationships between reading nodes in each time-window are presented in Figures 5, 6, 7, and 8. Table 2 provides the details of posterior means of each connection in each time-window. Because the main aim was to examine the inter-regional causal connections in the reading network, the connections among bilateral IFG, vOT and OCC were the focus. Self-connections (e.g., LvOT-to-LvOT) and horizontal lateral connections (e.g., LvOT-to-RvOT) are not provided here (but see

supplementary table 2 for the details). As stated in the Introduction section, RW vs. PW, PW vs. CS and CS vs. FF were estimated to investigate the roles of different linguistic properties in the early feedback from left inferior frontal gyrus to left ventral occipito-temporal cortex; RW vs. FF was used to replicate the finding reported in Woodhead et al. (2014). The DCM results for each comparison are reported below.

Table 2.

Posterior means for feedforward and feedback connections in different comparisons. Values in bold were significantly stronger for the former (> 0) or weaker for the latter (< 0) than the threshold (90 %) (Corresponding p values were smaller than 0.05).

	RW vs. FF (overall)			RW vs. PW (lexicality effect)		
	1-100 ms	1-150 ms	1-200 ms	1-100 ms	1-150 ms	1-200 ms
<i>Feedforward</i>						
LvOT to LIFG	-0.071	0.103	-0.050	-0.011	0.011	0.048
LOCC to LvOT	-0.190	-0.275	0.169	-0.123	-0.120	-0.025
LOCC to LIFG	-0.017	-0.126	-0.009	0.119	0.103	-0.102
RvOT to RIFG	-0.028	0.057	-0.002	-0.011	-0.088	0.054
ROCC to RvOT	0.093	-0.053	-0.022	-0.209	-0.031	-0.154
ROCC to RIFG	-0.245	-0.029	-0.029	-0.115	-0.138	-0.054
<i>Feedback</i>						
LIFG to LvOT	-0.039	0.059	0.021	0.018	0.073	0.102
LvOT to LOCC	-0.026	-0.154	-0.132	-0.010	0.029	-0.195
LIFG to LOCC	-0.020	-0.018	-0.006	0.006	0.022	0.004
RIFG to RvOT	0.006	0.035	-0.023	-0.019	0.008	-0.010
RvOT to ROCC	-0.042	-0.018	-0.038	0.034	0.048	0.064
RIFG to ROCC	-0.037	-0.009	0.047	0.004	0.066	-0.012

	PW vs. CS (phonological effect)			CS vs. FF (letter effect)		
	1-100 ms	1-150 ms	1-200 ms	1-100 ms	1-150 ms	1-200 ms
<i>Feedforward</i>						
LvOT to LIFG	0.036	-0.024	-0.009	0.029	0.038	0.133
LOCC to LvOT	-0.254	-0.115	0.066	0.124	-0.018	-0.058
LOCC to LIFG	0.098	-0.008	0.051	-0.192	0.075	-0.045
RvOT to RIFG	0.020	0.003	0.032	0.044	0.070	-0.016
ROCC to RvOT	-0.034	0.085	0.155	0.317	0.087	0.080
ROCC to RIFG	0.073	0.195	-0.095	0.090	0.077	0.015
<i>Feedback</i>						
LIFG to LvOT	-0.016	-0.017	0.068	-0.027	-0.120	-0.106
LvOT to LOCC	-0.014	0.003	-0.006	-0.046	-0.042	0.027
LIFG to LOCC	0.016	0.014	0.055	-0.003	0.036	-0.021
RIFG to RvOT	-0.007	0.058	0.037	-0.063	0.078	0.155
RvOT to ROCC	0.028	0.038	-0.037	-0.065	-0.051	-0.011
RIFG to ROCC	0.005	0.012	-0.030	0.001	-0.008	-0.028

RW vs. FF comparison (overall). As shown in Figure 5 and Table 2, the main results of the RW vs. FF comparison are as follows. (1) The left hemisphere (LH) feedback connections being significantly stronger for one condition than another condition did not occur in the first 100 ms but occurred in the first 150 ms time-window. However, the LvOT-to-LIFG and LOCC-to-LvOT feedforward connections were apparent in the 1-100 ms time-window and both of them were stronger for FF than for RW. (2) The LIFG-to-LvOT feedback and LvOT-to-LIFG feedforward connections were stronger for RW than for FF in the 1-150 ms time-window. (3) Significant connections mainly occurred in the LH; in the right hemisphere (RH) no significant feedback or feedforward connections were found in the 1-150 ms and 1-200 ms time-windows.

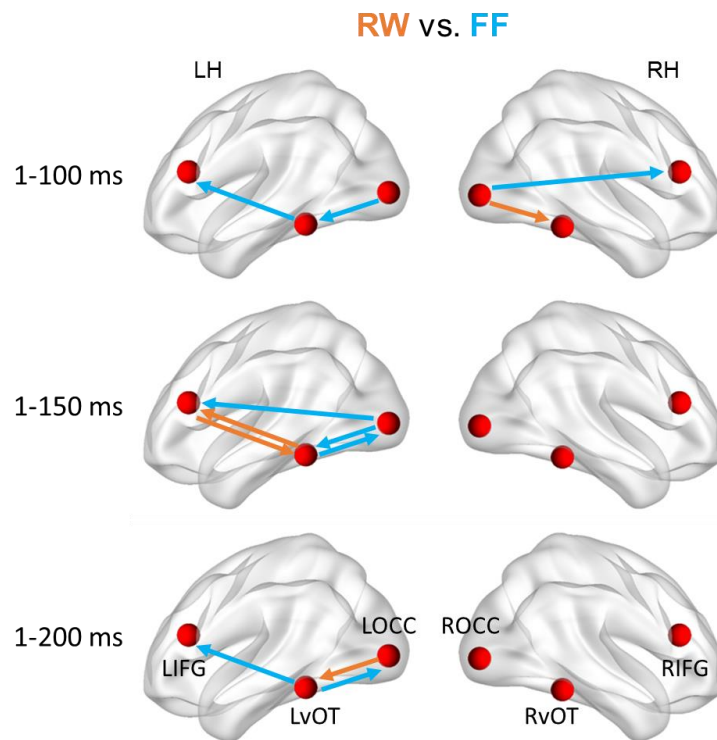


Figure 5. RW vs. FF comparison for inter-regional causal connections in three time-windows, 1-100 ms, 1-150 ms, and 1-200 ms. Only significant causal connections are shown here. The arrows denote directions of causal influences. An orange line denotes a connection that is stronger for RW than for FF while a blue line denotes a connection that is stronger for FF than for RW.

RW vs. PW comparison (Lexicality effect). As shown in Figure 6 and Table 2, the main results of the RW vs. PW comparison are as follows. (1) The LIFG-to-LvOT feedback connection was stronger for RW than for PW in both the 1-150 ms and 1-200 ms time-windows. (2) The feedback or feedforward connections between LOCC and LvOT were stronger for PW than for RW in all time-windows. (3) The feedforward connections in the RH were all stronger for PW than for RW; in the meantime, no significant feedback connections were found in the RH. (4) No significant feedback connections were found in the 1-100 ms time-window. (5) The significant feedforward connections were mainly found to be the connections from OCC to IFG, and from OCC to vOT in both hemispheres in all windows.

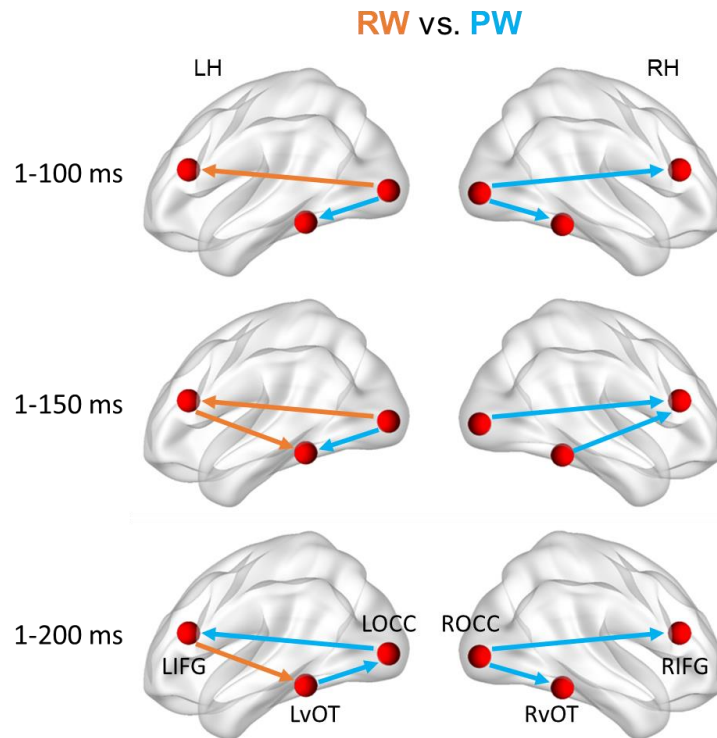


Figure 6. RW vs. PW comparison (lexicality effect) for inter-regional causal connections in three time-windows, 1-100 ms, 1-150 ms, and 1-200 ms. Only significant causal connections are shown here. The arrows denote directions of causal influences. An orange line denotes a connection that is stronger for RW than for PW while a blue line denotes a connection that is stronger for PW than for RW.

PW vs. CS comparison (Phonological effect). As shown in Figure 7 and Table 2, the main results of the PW vs. CS comparison are as follows. (1) The LIFG-to-LvOT feedback connection was stronger for PW than for CS in the 1-200 ms time-windows, which was accompanied by the LOCC-to-LvOT and LOCC-to-LIFG feedforward connections being stronger for PW than for CS. (2) The LOCC-to-LvOT feedforward connection was stronger for CS than for PW in both 1-100 ms and 1-150 ms time-windows. (3) All RH connections except the ROCC-to-RIFG connection were stronger for PW than for CS in all three time-windows. (4) No significant feedback connection was found in the 1-100 ms time-window. (5) All significant feedforward connections were found to be the connections from OCC to IFG, and from OCC to vOT in both hemispheres.

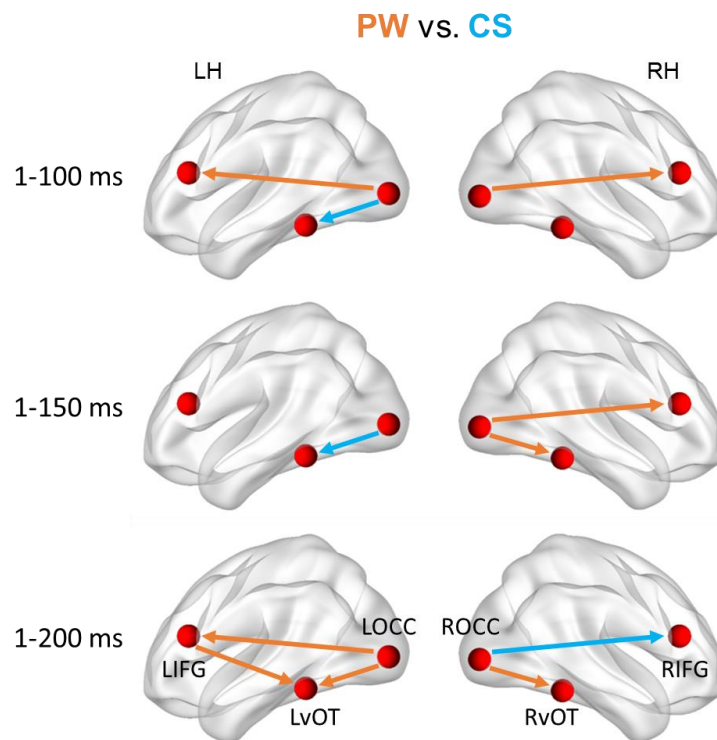


Figure 7. PW vs. CS comparison (phonological effect) for inter-regional causal connections in three time-windows, 1-100 ms, 1-150 ms, and 1-200 ms. Only significant causal connections are shown here. The arrows denote directions of causal influences. An orange line denotes a connection that is stronger for PW than for CS while a blue line denotes a connection that is stronger for CS than for PW.

CS vs. FF comparison (Letter effect). Details of the CS vs. FF comparison are provided in Figure 8 and Table 2. The main results are as follows. (1) The LIFG-to-LvOT feedback connection was stronger for FF than for CS in both 1-150 ms and 1-200 ms time-windows, which was accompanied by the RIFG-to-RvOT connection being stronger for CS than for FF in both time windows. (2) All significant feedforward connections in the RH were stronger for CS than for FF. (3) No significant feedback connection was found in the 1-100 ms time-window. (4) The significant feedforward connections were mainly found to be the connections from OCC to IFG, and from OCC to vOT in both hemispheres.

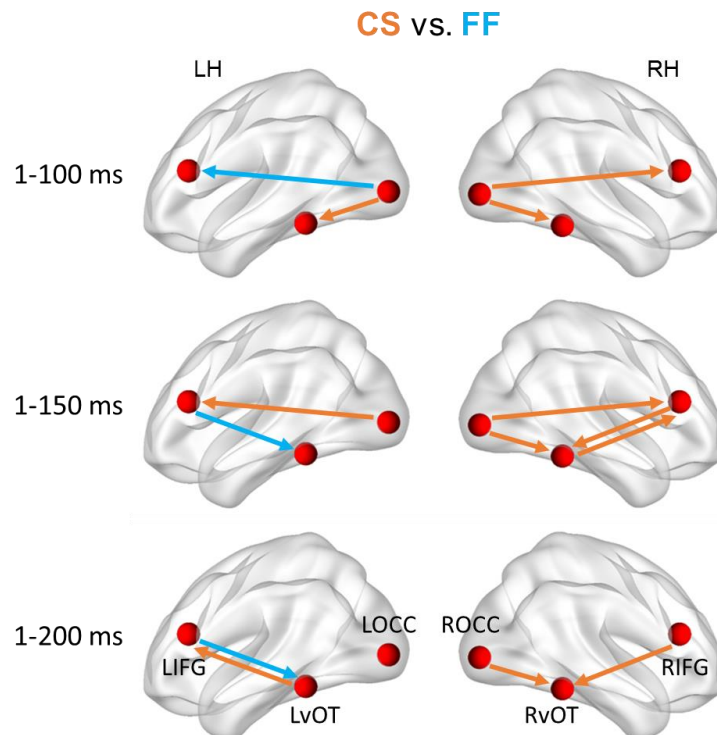


Figure 8. CS vs. FF comparison (letter effect) for inter-regional causal connections in three time-windows, 1-100 ms, 1-150 ms, and 1-200 ms. Only significant causal connections are shown here. The arrows denote directions of causal influences. An orange line denotes a connection that is stronger for CS than for FF while a blue line denotes a connection that is stronger for FF than for CS.

Discussion

With the application of the DCM technique to MEG data, Woodhead et al. (2014) previously investigated top-down feedback from LIFG to LvOT at the early stages of visual word recognition. Their results showed that the LIFG-to-LvOT feedback connection was stronger for real words than for false font stimuli within the first 200 ms after stimulus onset. Following this study, the current study sought to examine the nature of this early top-down feedback from LIFG to LvOT. Four types of visual stimulus, real words, pseudowords, consonant strings, and meaningless false fonts were included. Three comparisons - real word vs. pseudowords, pseudowords vs. consonant strings, and consonant strings vs. false fonts - were designed to decompose what linguistic properties of visual words trigger this early top-down feedback.

Overall, the findings were that the LIFG-to-LvOT feedback was stronger for real words than for false fonts in the 1-150 ms time-window, replicating the basic observation in Woodhead et al. (2014). The current study further found that this top-down feedback was stronger for real words than for pseudowords in both 1-150 and 1-200 ms time-windows and was stronger for pseudowords than for consonant strings in the 1-200 ms time-window, indicating that compared to phonological properties, lexical-semantic properties have an earlier influence on the top-down feedback from LIFG to LvOT; but the RIFG-to-RvOT feedback did not show any differences between conditions, indicating a left-lateralised network for the processing of high-level properties of words. In addition, the LIFG-to-LvOT feedback was stronger for false fonts than for consonant strings while the RIFG-to-RvOT feedback was stronger for consonant strings than for false fonts. The findings of the current study are broadly in line with the Interactive Account of LvOT in word recognition which proposes that the LvOT is an interface linking bottom-up visual inputs from occipital cortex and top-down linguistic

influences from language areas (Price & Devlin, 2011), and have implications for better understanding the neural correlates of reading, especially the early neural dynamics between reading-related areas.

Early Feedback from Frontal to Ventral Occipito-Temporal Cortex

The real words vs. false fonts comparison revealed that the LIFG-to-LvOT top-down feedback was stronger for real words than for false fonts in the 1-150 ms time-window. This finding is in line with the findings reported in Woodhead et al. (2014) which found that this top-down feedback was stronger for words than for false fonts in the 1-200 ms time-window. In their MEG study, a semantic categorisation task (whether a stimulus on the screen is a personal name, e.g., *john*) was adopted and six-node network was used in the DCM analysis, which is similar to the present study. However, that study used 1-100 ms, 1-200 ms, and 1-300 ms time-windows, but didn't use 1-150 ms time-window. The present finding of the LIFG-to-LvOT feedback in the 1-150 ms time-window thus extends Woodhead et al (2014) to suggest that the top-down feedback for words occurs in an earlier time-window. Because the current study and Woodhead et al. (2014) both found this early top-down feedback for words, we argue that the early top-down feedback from the anterior higher-order frontal cortex to the posterior lower-order ventral occipito-temporal cortex is reliable and robust. Both studies suggest that word representations encoded in the LIFG send rapid feedback predictions to LvOT where visual sensory inputs converge such that visual words can be rapidly recognised. Similar to Woodhead et al. (2014), our experimental design also removed any influences of participants' expectations in the early top-down feedback. That is, the mixture of four different trial types and the randomisation of stimulus presentation order in the present study meant that participants could not form any expectations about upcoming stimuli, removing to a large degree the possibility of subject prediction and expectation. Interestingly, we also observed

that the LvOT-to-LOCC feedback connections were stronger for real words than for false fonts. It is likely that as false fonts are unfamiliar, they recruited more low-level feature processing, and thus need more neural interactions between these two visual areas.

To gain further insight into the nature of the LIFG-to-LvOT top-down feedback, we used three different comparisons, real words vs. pseudowords to index lexicality effects, pseudowords vs. consonant strings to index phonological effects, and consonant strings vs. false fonts to index letter effects. The DCM results showed that in the 1-200 ms time-window, the LIFG-to-LvOT feedback was stronger for real words than for pseudowords, indicating lexical-semantic influence; and it was stronger for pseudowords than for consonant strings, indicating phonological influence. Interestingly, this connection was also stronger for false fonts than for consonant strings, indicating the influences of letter combinations. These results collectively indicate that lexical-semantic, phonological and letter properties all rapidly drive this early top-down feedback at the very early stages. By decomposing the real words vs. false fonts comparison into different comparisons reflecting lexical-semantic, phonological and letter effects, the current study allows us to extend the previous proposal that this top-down feedback may reflect phonological constraints on visual processing (Woodhead et al., 2014).

The lexical-semantic, phonological and letter effects in early top-down feedback are consistent with the existing literature. The finding that lexical-semantic properties engaged in processing real words can begin influencing visual word recognition before 200 ms after stimulus onset is consistent with data from previous EEG/MEG studies (e.g., Assadollahi & Pulvermuller, 2001, 2003; Sereno et al., 1998; Hauk et al., 2006, 2012; Mahé et al., 2012; Araújo et al., 2015). Hauk et al. (2006) conducted a linear regression analysis in EEG data and found an early difference between words and pseudowords at around 160 ms (Hauk et al., 2006). In a lexical decision task, Braun, Hutzler, Ziegler, Dambacher, and Jacobs (2009) found the earliest differences between pseudohomophones (e.g., *ROZE*) and matched spelling

controls (e.g., *ROFE*) at around 150 ms after stimulus onset, indicating the involvement of phonological information at the early stages of visual word recognition. The neural differences between pseudowords and consonant strings have also been observed within 200 ms of stimulus onset (e.g., Coch & Mitra, 2015; Araújo et al., 2015). The influence of letter information on the top-down feedback was out of our expectation because the LIFG is usually thought to be an area mainly for high-level lexical-semantic and phonological processing (e.g., Taylor et al., 2013; Carreiras et al., 2014) and not for letter processing. However, some studies have shown that the LIFG involves letter processing (e.g., Flowers et al., 2004; Liu et al., 2010, 2011; Vartiainen et al., 2011), and letters and non-letters can be discriminated at around 150 ms after stimulus onset (e.g., Bentin et al., 1999; Appelbaum et al., 2009; Mahé et al., 2012; Araújo et al., 2015; Coch & Meade, 2016). The previous observations presented here suggest influences of lexical-semantic, phonological and low-level letter information can all appear in the first 200 ms, which is consistent with the current finding that lexical-semantic, phonological and low-level letter information potentially rapidly affect the early feedback from frontal cortex to ventral occipito-temporal cortex.

It was further found in this study that the RIFG-to-RvOT feedback connection did not show differences between two conditions in the comparisons of real words vs. pseudowords, pseudowords vs. consonant strings, and real words vs. false fonts. The absence of the RIFG-to-RvOT feedback connection suggests that even at the early stages of visual word recognition, the top-down feedback from frontal cortex to ventral occipito-temporal cortex is already strongly left-lateralised. This replicates exactly Woodhead et al.'s (2014) observation that the LIFG-to-LvOT feedback was stronger for words than for false fonts before 200 ms but the RIFG-to-RvOT feedback was absent. The left-lateralised IFG-to-vOT feedback connection pattern found here probably reflects the top-down influences of the left hemisphere spoken language network on the LvOT in reading. Cai et al. (2008) previously observed that the

activity in the posterior occipito-temporal cortex involved in visual word recognition is co-lateralised to the same hemisphere as language production, indicating the top-down influences of frontal language area in the development of posterior occipito-temporal cortex. The finding here that the top-down feedback was highly left-lateralised is consistent with the observation in that study. Previous MEG studies have also observed a left-lateralised activation in the inferior frontal cortex in the first 200 ms during reading-related tasks (e.g., Pammer et al., 2004; Cornelissen et al., 2009; Wheat et al., 2010; Klein et al., 2014).

Earlier LIFG-to-LvOT Feedback Driven by Lexical-Semantic Properties

The different comparisons between conditions across different time-windows can reveal which effect commences earlier or later in driving the LIFG-to-LvOT feedback after the presentation of written words. It was clearly seen that the LIFG-to-LvOT connection was stronger for real words than for pseudowords in both the 1-150 ms and 1-200 ms time-windows, and this same connection was in turn stronger for pseudowords than for consonant strings in the later 1-200 ms time-window. This probably indicates that, although both lexical and phonological factors can trigger rapid LIFG-to-LvOT feedback, the lexical-semantic properties of a familiar word start influencing word recognition processes earlier than its phonological elements.

Two potential factors could explain why the stronger LIFG-to-LvOT connection for real words than pseudowords occurred in the early 1-150 time-window but the stronger LIFG-to-LvOT connection for pseudowords than consonant strings occurred in the late 1-200 time-window. Firstly, cognitive models of single word reading (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) propose that during single word comprehension, the semantic system can be accessed either via direct route from orthography or indirectly via phonological mediation.

When an orthographic entry has a corresponding meaning (i.e., when it is a real word), it could be more efficient for a skilled reader to access meaning via direct route from orthography to semantics (Taft & van Graan, 1998). Taft and van Graan (1998) found that in skilled readers, differences in reaction time between regular words and irregular words were observed in a naming task which is more phonologically based; however, these differences were washed out in a semantic categorisation task which is more semantic-based. This indicates that accessing semantics can be realised directly from orthographic strings when semantic information is explicitly required by task. Participants in our study were all adult readers with extensive reading experience, so the lexical-semantic information could be extracted directly from orthography and rapidly trigger a top-down feedback from frontal cortex to help word form analysis at early stages. That may result in the stronger LIFG-to-LvOT feedback for words than for pseudowords in the early 1-150 ms window. Because the task goal is to judge whether a word is an animal word after knowing its meaning, this top-down feedback continues to exist in the 1-200 ms window. In contrast, constructing phonological information via a phonologically-mediated route is necessary for pseudowords and there is also no corresponding semantic information, which could collectively slow down the influences on the feedback from frontal cortex to ventral occipito-temporal area, as can be seen in the pseudowords vs. consonant strings comparison where the LIFG-to-LvOT feedback was stronger for pseudowords than for consonant strings in the late 1-200 ms but not the early 1-150 ms windows.

Secondly, it has been shown that lexical-semantic properties of visual words exert an influence on visual word recognition at around 150 ms after stimulus onset. Assadollahi and Pulvermuller (2001, 2003) found that word frequency effect is observed at around 150 ms post stimulus onset. Sereno et al. (1998) found that the difference between real words and pseudowords occurred at around 112 ms and word frequency effect at around 144 ms. The

difference between words and pseudowords was also observed at around 160 ms in the studies with lexico-semantic emphasis such as lexical decision or semantic categorisation tasks (e.g., Hauk et al., 2006, 2012; Mahé et al., 2012; Araújo et al., 2015; Coch & Meade, 2016). It is likely that the earliness of the neural responses to lexical properties makes it possible for the early LIFG-to-LvOT top-down feedback to occur.

Simultaneity of Letter and Lexicality Effects

Interestingly, the RIFG-to-RvOT feedback connection was stronger for consonant strings than false fonts in both the 1-150 and 1-200 ms time-windows, indicating that this feedback connection in the right hemisphere may play a role in the processing of low-level surface properties in letter combinations (Bentin et al., 1999; Appelbaum et al., 2009). At the same time, this right hemisphere feedback was accompanied by the LIFG-to-LvOT feedback connection that was stronger for false fonts than consonant strings in the same two time-windows. The earliness of the significant IFG-to-vOT connections for consonant strings or false fonts may reflect an early differentiation between letters and non-letters. This interpretation is consistent with the observation that letters and non-letters can be discriminated around 150 ms after stimulus onset (e.g., Bentin et al., 1999; Appelbaum et al., 2009; Coch & Meade, 2016). In an EEG study, Bentin et al. (1999) observed an early N170 difference between false fonts and letter strings in the bilateral occipito-temporal sites, which was further confirmed in other EEG studies (Appelbaum et al., 2009; Coch & Meade, 2016). These findings indicate that human brain can rapidly detect the differences between letters and non-letters. The current study further found that the brain uses inferior frontal cortex to modulate vOT during the process of letter detection.

The letter effect on the top-down feedback occurred in the 1-150 ms and 1-200 ms time-windows. Interestingly, in the same time-windows, the LIFG-to-LvOT connection was

stronger for real words than for pseudowords, and was also stronger for real words than for false fonts, indicating a lexicality effect. In other words, the early letter-specific effects and lexicality-specific effects in the top-down feedback from IFG to vOT appeared simultaneously; specifically, when rapid letter detection triggers a bilateral top-down feedback from IFG to vOT, high-level lexical properties also rapidly trigger a left-lateralised top-down feedback from IFG to vOT. The bilateral top-down feedback from IFG to vOT within 200 ms after stimulus onset is generally consistent with previous findings that the early neural difference between false fonts and consonant strings occurs bilaterally (e.g., Bentin et al., 1999; Coch & Meade, 2016). The combination of the left-lateralised top-down feedback for real words in both the 1-150 ms and 1-200 windows, and for pseudowords in the 1-200 ms window may reflect the influences of spoken language areas in the development of the posterior vOT in reading (e.g., Cai et al., 2008; Hannagan et al., 2015).

It was also found that significant feedforward connections predominantly occurred in the connections from occipital cortex to vOT cortex and from vOT cortex to IFG, while significant feedback connections predominantly occurred in the connections from IFG to vOT cortex. This suggests that different connections between OCC, vOT and IFG play different roles in visual word recognition. It could be that the IFG-to-vOT connection is mainly for transferring influences from high-level linguistic properties stored in IFG to orthographic form analysis in vOT; in the meantime, the OCC-to-vOT and vOT-to-IFG connections are mainly for transferring sensory inputs from occipital cortex to ventral visual pathway and inferior frontal cortex. However, this possibility still lacks evidence. In addition, the significant LOCC-to-LvOT and LvOT-to-LIFG feedforward connections occurred in the 1-100 ms time-window, which was earlier than the significant LIFG-to-LvOT feedback connection which occurred in the 1-150 and 1-200 ms time-windows. This may indicate that relevant information on visual word recognition is firstly processed in a pure low-high hierarchical manner, i.e., information

flows from low-level to high-level areas (LOCC-to-LvOT and -LIFG), followed by a combined high-low and low-high hierarchical manner, i.e., information flows from low-level to high-level areas coupled by information flows from high-level to low-level areas (LIFG-to-LvOT). However, this pattern was not observed in the real word vs. false fonts comparison in Woodhead et al. (2004) which found that the significant LIFG-to-LvOT feedback connection was earlier than all significant feedforward connections. Note that both their study and the current study examined the differences in connectivity between two conditions, thus a non-significant connection possibly means that the connection in both conditions exists but the strength is different. However, the finding still suggests that neural information flow specific for a linguistic process (e.g., lexical-semantic, phonological or letter processing) follows that top-down information follows bottom-up information.

Limitations

The present study had several limitations. The first one is that the number of participants recruited in the current study was 15, which is a small sample. This may decrease the robustness of our results. In this sense, a larger sample would be preferable. The second one is that a template brain, but not individual anatomical brains, was used in the DCM analysis. While this acceptable approach is widely used in the field, the use of individual anatomical brains may improve the accuracy of source reconstruction and further improve the accuracy of the computation of inter-regional effective connectivity. The third one is the small proportion of catch trials (i.e., animal words) in all stimuli, which could affect the top-down feedback. The ratio of catch trials to critical trials was 1:5, and participants only needed to respond to animal words but not to critical trials. This approach with a small ratio of catch trials to critical trials has been widely used in the M/EEG studies (Woodhead et al., 2014; Coch, 2015; Coch & Meade, 2016; Strijkers et al., 2015). Evidence from studies on response inhibition using the

GO/NO-GO paradigm suggest that the ratio of GO trials to NO-GO trials strongly affects early neural responses to stimuli (e.g., Nieuwenhuis et al., 2003). In a general sense, the finding indicates that the frequency of response such as button press highly affects early neural activity. Thus, an open question is how different ratios affect the early top-down feedback in this study. A more specific question is whether different ratios change the influences of lexico-semantic processing in the early top-down feedback in a semantic task, and how if so. Future studies could examine this issue. In addition, the current study designed only 32 models to reduce the number of models due to the limited computation, which possibly led a bias in the estimation of information flow in the word-recognition network. A recent study applied cloud-computation to the estimation of more than 4000 models to establish bias-free information flow maps in the brain (Ge et al., 2015). Future investigations could adopt this approach to avoid a bias in designing models of the word-recognition network.

Conclusion

With the application of DCM techniques to MEG data, the current study aimed to examine the nature of early top-down feedback from frontal to ventral occipito-temporal cortex during visual word recognition. The finding that the LIFG-to-LvOT connection was stronger for real words than for false fonts in the 1-150 ms time-window replicates the observation in Woodhead et al. (2014). More importantly, we established that lexical-semantic, phonological, and letter information all could influence the early LIFG-to-LvOT top-down feedback; lexical-semantic and letter information both occur rapidly and simultaneously, and earlier than phonological information. The findings are in close alignment with the Interactive Account of LvOT (Price & Devlin, 2011), and are important in providing more details about the nature of early top-down feedback during visual word recognition observed in Woodhead et al. (2014).

Appendix A**Supplementary table 1.**

List of eight MEG/EEG studies in the current ALE meta-analysis.

Article	Tools	Tasks	Sample	Age (years)	MNI coordinates (x y z)
<i>Pammer et al., 2004</i>	MEG	lexical decision	10	34	-14 -88 -6 30 -94 -6 -32 -64 -6 -60 8 22
<i>Cornelissen et al., 2009</i>	MEG	passive viewing	10	N/A	-36 -84 -2 -46 -56 -18 -50 26 8
<i>Wheat et al., 2010</i>	MEG	word reading in a masked pseudohomophone priming task	20	23.2	-26 -96 8 -56 4 18
<i>Yvert et al., 2012</i>	EEG	phoneme deletion and semantic categorisation	15	24.5	-42 -84 -5 42 -84 -5 -62 -40 2 62 -40 2 -62 -52 -12 62 -52 -12 -54 0 -20 54 0 -20
<i>Thesen et al., 2012</i>	MEG	semantic categorisation	12	23	-40 -78 -18 -46 -52 -20
<i>Klein et al., 2014</i>	MEG	word naming and lexical decision in a masked onset priming task	48	20.9	-58 8 22 -26 -94 5
<i>Quinn et al., 2014</i>	MEG	semantic categorisation	5	24.44	-30 -96 2 -44 -58 -14 -44 -2 24
<i>Woodhead et al., 2014</i>	MEG	semantic categorisation	10	57	-21 -90 8 -43 -60 -16 -44 23 -7

Appendix B

Supplementary Table 2.

Posterior means for feedforward, feedback, lateral, and self- connections in different comparisons. Values in bold were significantly stronger for the formers (> 0) or weaker for the latters (< 0) than the chance (90 %) (Corresponding p values were smaller than 0.05).

Feedback	RW vs. FF					
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LIFG to LvOT	-0.039	0.061	0.059	0.046	0.021	0.047
LvOT to LOCC	-0.026	0.086	-0.154	0.069	-0.132	0.062
LIFG to LOCC	-0.020	0.072	-0.018	0.053	-0.006	0.046
RIFG to RvOT	0.006	0.056	0.035	0.050	-0.023	0.044
RvOT to ROCC	-0.042	0.087	-0.018	0.071	-0.038	0.056
RIFG to ROCC	-0.037	0.073	-0.009	0.055	0.047	0.046
<i>Feedforward</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LvOT to LIFG	-0.071	0.054	0.103	0.054	-0.050	0.039
LOCC to LvOT	-0.190	0.055	-0.275	0.053	0.169	0.048
LOCC to LIFG	-0.017	0.046	-0.126	0.041	-0.009	0.027
RvOT to RIFG	-0.028	0.052	0.057	0.056	-0.002	0.043
ROCC to RvOT	0.093	0.057	-0.053	0.056	-0.022	0.051
ROCC to RIFG	-0.245	0.050	-0.029	0.041	-0.029	0.026
<i>Lateral</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
ROCC to LOCC	0.116	0.080	0.101	0.066	-0.005	0.0584
RvOT to LvOT	0.032	0.085	0.044	0.072	-0.116	0.0694
RIFG to LIFG	-0.039	0.065	-0.035	0.051	0.0436	0.0518
LOCC to ROCC	-0.020	0.081	0.015	0.064	-0.077	0.0586
LvOT to RvOT	0.068	0.083	0.144	0.069	0.0632	0.0665
LIFG to RIFG	-0.007	0.067	-0.072	0.055	0.0932	0.0516
<i>Self-connections</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LOCC to LOCC	0.123	0.018	0.030	0.015	-0.016	0.0101
ROCC to ROCC	0.055	0.016	0.014	0.017	-0.041	0.0087
LvOT to LvOT	0.018	0.040	0.061	0.026	-0.066	0.0212
RvOT to RvOT	0.131	0.039	-0.022	0.029	-0.035	0.021
LIFG to LIFG	0.110	0.038	0.026	0.025	0.0035	0.0179
RIFG to RIFG	0.160	0.037	0.003	0.025	-0.047	0.0206

(continued)

Supplementary Table 2. (continued)

<i>Feedback</i>	RW vs. PW (lexicity effect)					
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LIFG to LvOT	0.018	0.042	0.073	0.057	0.102	0.043
LvOT to LOCC	-0.010	0.088	0.029	0.069	-0.195	0.067
LIFG to LOCC	0.006	0.078	0.022	0.040	0.004	0.041
RIFG to RvOT	-0.019	0.036	0.008	0.056	-0.010	0.043
RvOT to ROCC	0.034	0.087	0.048	0.072	0.064	0.059
RIFG to ROCC	0.004	0.070	0.066	0.038	-0.012	0.048
<i>Feedforward</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LvOT to LIFG	-0.011	0.071	0.011	0.068	0.048	0.056
LOCC to LvOT	-0.123	0.067	-0.120	0.049	-0.025	0.044
LOCC to LIFG	0.119	0.051	0.103	0.035	-0.102	0.036
RvOT to RIFG	-0.011	0.063	-0.088	0.062	0.054	0.059
ROCC to RvOT	-0.209	0.063	-0.031	0.050	-0.154	0.048
ROCC to RIFG	-0.115	0.049	-0.138	0.042	-0.054	0.037
<i>Lateral</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
ROCC to LOCC	0.085	0.079	-0.036	0.067	0.0122	0.0638
RvOT to LvOT	0.082	0.082	0.047	0.077	-0.043	0.0704
RIFG to LIFG	0.016	0.057	-0.074	0.053	-0.071	0.0534
LOCC to ROCC	0.031	0.079	-0.114	0.075	-0.134	0.0624
LvOT to RvOT	-0.013	0.082	0.068	0.077	-0.03	0.0689
LIFG to RIFG	-0.010	0.059	-0.057	0.052	-0.057	0.053
<i>Self-connections</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LOCC to LOCC	-0.135	0.023	0.005	0.012	0.0545	0.0084
ROCC to ROCC	0.093	0.019	0.111	0.014	0.0837	0.0108
LvOT to LvOT	0.051	0.044	-0.028	0.024	0.0001	0.0228
RvOT to RvOT	0.176	0.045	-0.038	0.022	-4E-04	0.0231
LIFG to LIFG	-0.056	0.047	-0.032	0.020	0.0417	0.0194
RIFG to RIFG	0.084	0.040	-0.004	0.021	0.0316	0.0239

(continued)

Supplementary Table 2. (continued)

<i>Feedback</i>	PW vs. CS (Phonological effect)					
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LIFG to LvOT	-0.016	0.050	-0.017	0.044	0.068	0.045
LvOT to LOCC	-0.014	0.083	0.003	0.073	-0.006	0.063
LIFG to LOCC	0.016	0.068	0.014	0.059	0.055	0.053
RIFG to RvOT	-0.007	0.048	0.058	0.046	0.037	0.044
RvOT to ROCC	0.028	0.081	0.038	0.073	-0.037	0.056
RIFG to ROCC	0.005	0.068	0.012	0.058	-0.030	0.050
<i>Feedforward</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LvOT to LIFG	0.036	0.062	-0.024	0.046	-0.009	0.045
LOCC to LvOT	-0.254	0.064	-0.115	0.051	0.066	0.052
LOCC to LIFG	0.098	0.047	-0.008	0.043	0.051	0.035
RvOT to RIFG	0.020	0.064	0.003	0.047	0.032	0.049
ROCC to RvOT	-0.034	0.060	0.085	0.052	0.155	0.049
ROCC to RIFG	0.073	0.047	0.195	0.040	-0.095	0.038
<i>Lateral</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
ROCC to LOCC	0.152	0.077	0.051	0.068	-0.18	0.067
RvOT to LvOT	-0.063	0.081	0.005	0.074	-0.02	0.07
RIFG to LIFG	0.016	0.058	-0.025	0.064	0.03	0.05
LOCC to ROCC	0.005	0.082	-0.047	0.069	-0.1	0.063
LvOT to RvOT	-0.048	0.083	0.040	0.075	0.109	0.072
LIFG to RIFG	-0.013	0.058	-0.031	0.062	0.092	0.05
<i>Self-connections</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LOCC to LOCC	0.027	0.021	-0.030	0.011	0.01	0.009
ROCC to ROCC	0.027	0.020	-0.012	0.010	-0.02	0.01
LvOT to LvOT	0.103	0.051	0.028	0.026	-0.01	0.022
RvOT to RvOT	-0.019	0.044	-0.076	0.024	-0.07	0.023
LIFG to LIFG	-0.104	0.049	-0.020	0.020	-0.07	0.019
RIFG to RIFG	-0.020	0.049	-0.044	0.021	0.019	0.022

(continued)

Supplementary Table 2. (continued)

<i>Feedback</i>	CS vs. FF (Low-level letter effect)					
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LIFG to LvOT	-0.027	0.069	-0.120	0.054	-0.106	0.053
LvOT to LOCC	-0.046	0.082	-0.042	0.073	0.027	0.061
LIFG to LOCC	-0.003	0.045	0.036	0.053	-0.021	0.039
RIFG to RvOT	-0.063	0.074	0.078	0.049	0.155	0.047
RvOT to ROCC	-0.065	0.081	-0.051	0.072	-0.011	0.063
RIFG to ROCC	0.001	0.045	-0.008	0.052	-0.028	0.046
<i>Feedforward</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LvOT to LIFG	0.029	0.062	0.038	0.057	0.133	0.049
LOCC to LvOT	0.124	0.059	-0.018	0.053	-0.058	0.053
LOCC to LIFG	-0.192	0.045	0.075	0.033	-0.045	0.039
RvOT to RIFG	0.044	0.064	0.070	0.053	-0.016	0.048
ROCC to RvOT	0.317	0.064	0.087	0.053	0.080	0.058
ROCC to RIFG	0.090	0.043	0.077	0.036	0.015	0.038
<i>Lateral</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
ROCC to LOCC	0.082	0.080	0.095	0.070	0.208	0.057
RvOT to LvOT	-0.059	0.083	-0.002	0.077	-0.07	0.074
RIFG to LIFG	-0.041	0.049	-0.032	0.042	0.097	0.055
LOCC to ROCC	0.032	0.077	0.056	0.065	0.171	0.055
LvOT to RvOT	0.032	0.082	0.008	0.075	0.13	0.067
LIFG to RIFG	-0.012	0.049	-0.005	0.045	-0.08	0.06
<i>Self-connections</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LOCC to LOCC	0.048	0.022	-0.087	0.015	-0	0.01
ROCC to ROCC	-0.071	0.018	-0.028	0.010	-0.1	0.012
LvOT to LvOT	-0.003	0.045	0.140	0.025	-0.03	0.025
RvOT to RvOT	-0.095	0.049	-0.017	0.026	-0.1	0.027
LIFG to LIFG	-0.030	0.042	-0.013	0.021	0.016	0.019
RIFG to RIFG	-0.128	0.046	-0.033	0.023	0.007	0.021

Chapter 3

Task modulation of early top-down feedback from frontal to ventral occipito-temporal cortex during visual word recognition

Yu Li, Sachiko Kinoshita, Paul Sowman, & Anne Castles

ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney

Abstract

The human brain is flexible in responding to the external environment. Neuroimaging experiments have revealed that the early neural responses to visual words are highly modulated by task goal. A previous study has found that the early top-down feedback from frontal cortex to ventral occipito-temporal cortex is stronger for words than for meaningless symbols within 200 ms after stimulus onset (Woodhead et al., 2014). Building on this, our previous MEG study with a semantic categorisation task further found that during visual word recognition, lexical-semantic properties rapidly triggered the top-down feedback from the left inferior frontal gyrus (LIFG) to the left ventral occipito-temporal cortex (LvOT), and their influences were earlier than phonological properties (Li et al., in preparation; Chapter 2). However, whether this early top-down feedback is dependent on task goals is still unclear. The present study aimed to investigate this issue by using a visual discrimination task which relies on low-level visual feature processing but not necessarily high-level linguistic properties. Fifteen young adults participated in the study. All four types of stimulus, real words, pseudowords, consonant strings, and false fonts, were the same as the previous study. The results showed that (1) the LIFG-to-LvOT connection was stronger for real words than for pseudowords in both 1-150 ms and 1-200 ms time-windows in the discrimination task, which is similar to that in the semantic task, indicating that the influences of lexical factors are task-independent; (2) the connection was stronger for consonant strings than for pseudowords in the 1-200 ms time-window and was stronger for false fonts than for consonant strings in the 1-150 ms time-window in the visual discrimination task. In contrast, the connection was stronger for pseudowords than for consonant strings in the 1-200 ms time-window and was stronger for consonant strings than for false fonts in the 1-150 ms time-window in the semantic categorisation task, indicating that the influences of phonological and letter-level properties are task-dependent. These findings are helpful for understanding how inter-regional directional neural connections during visual word recognition adapt to different task goals.

Introduction

Representing sound and meaning from writing symbols recruits and reshapes existing neural structures. There is mounting evidence that reading behaviour involves distributed brain regions with different functions, including the primary occipital cortex, the left ventral occipito-temporal cortex (LvOT) and the left inferior frontal gyrus (LIFG) (e.g., Turkeltaub, Eden, Jones, & Zeffiro, 2002; McCandliss, Cohen, & Dehaene, 2003; Bolger, Perfetti, & Schneider, 2005; Houde, Rosi, Lubin, & Joliot, 2010; Dehaene & Cohen, 2011; Price & Devlin, 2011; Cattinelli, Borghese, Gallucci, & Paulesu, 2013; Taylor, Rastle, & Davis, 2013). It is widely thought that the primary occipital cortex is responsible for visual processing, LvOT for orthographic form analysis and LIFG for high-level phonological and lexical-semantic processing in visual word recognition (see a recent meta-analysis of fMRI studies of reading, Taylor et al., 2013). These regions not only are consistently activated in relevant tasks (e.g., passive viewing, lexical decision and reading aloud tasks), but also are functionally connected to each other when recognising visual words (e.g., Horwitz, Rumsey, & Donohue, 1998; Pugh et al., 2000; van der Mark et al., 2011; Finn et al., 2013; Boets et al., 2013; Schurz et al., 2014) and at rest (e.g., Koyama et al., 2011; Vogel, Miezin, Peterson, & Schlaggar, 2012; Schurz et al., 2014; Li et al., 2017). A recent MEG study found that early top-down feedback from frontal cortex to ventral occipito-temporal cortex was stronger for words than for meaningless symbols within 200 ms after stimulus onset (Woodhead et al., 2014). With the inclusion of words, pseudowords, consonant strings, and false fonts, Li et al. (in preparation) further found that lexical-semantic properties and letter-level effects rapidly triggered the top-down feedback from LIFG to LvOT, and their influences were earlier than phonological properties (see Chapter 2). By applying dynamic causal modeling (DCM) to MEG data, the aim of the current study was to investigate whether task goal modulates this early top-down feedback from frontal cortex to ventral occipito-temporal cortex.

The extant literature has made great inroads into understanding the underlying cognitive mechanisms of visual word recognition (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Seidenberg & McClelland, 1989; Grainger & Ziegler, 2011; Norris, 2013). A range of cognitive processes including letter recognition, phonological decoding and lexical-semantic access are engaged during visual word recognition, and possible relationships between them are usually examined via a specific reading-related task. To date, many tasks have been used, including lexical decision, reading aloud, and semantic categorisation. The lexical decision task is probably the most frequently used task, in which participants are asked to decide whether a visual string is a real word, and to press one button if it is a real word and another button if it is not. This task is thought to recruit the main cognitive processes required in word recognition, including visual feature detection, letter binding, orthographic/phonological form analysis, and lexical-semantic access. It has been used to establish and refine models of reading and visual word recognition (e.g., Coltheart et al., 2001; Grainger & Ziegler, 2011) and to uncover relevant deficits in individuals with reading impairment (Vellutino, Fletcher, Snowling, & Scanlon, 2004; Ziegler & Goswami, 2005).

However, different tasks emphasise different cognitive processes involved in visual word recognition. Indeed, behavioural studies have shown that the influence of linguistic variables, such as word frequency, consistency between spelling and sound, and imageability, on visual word recognition are highly dependent on task goal (e.g., Monsell, Doyle, & Haggard, 1989; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Balota & Yap, 2006; Yap & Balota, 2009). For example, the word frequency effect - reflecting lexical-semantic access - contributes more to the performance of a lexical decision task than to a speeded naming task (Balota & Chumbley, 1984; Balota et al., 2004; Balota & Yap, 2006; Yap & Balota, 2009) and the word frequency effect observed in a lexical decision task has been found to be absent in same-different tasks (Norris & Kinoshita, 2008). These results suggest that the influence of

linguistic properties on visual word processing are not purely stimulus-driven but are strongly guided by task goals, with different emphases of cognitive resources in different tasks influencing behavioural outputs (e.g., Kinoshita & Norris, 2012).

Similarly, fMRI studies have also revealed that neural correlates of visual word processing, such as regional activation, are also modulated by task goals (e.g., Valdois et al., 2006; Carreiras, Mechelli, & Price, 2006; Carreiras, Mechelli, Estevez, & Price, 2007; Hon, Thompson, Sigala, & Duncan, 2009; Baker et al., 2007; Starrfelt & Gerlach, 2007; Wright et al., 2008; Twomey, Duncan, Price & Devlin, 2011; Yang, Wang, Shu, & Zevin, 2012; Mano et al., 2013; Yang & Zevin, 2014; Graves, Boukrina, Mattheiss, Alexander, & Baillet, 2016). For example, lexical decision tasks and semantic categorisation tasks appear to rely more on semantic systems in the brain (e.g., middle temporal gyrus and anterior temporal lobe) while naming and reading aloud tasks rely more on phonological systems (e.g., superior temporal gyrus and inferior frontal gyrus) (McNorgan, Chabal, O'Young, Lukic, & Booth, 2015). Effective connectivity can serve as a tool to examine the top-down causal influences of a brain area to another area. Studies using effective connectivity have also shown that the inter-regional directional influences during visual word processing are task-dependent (e.g., Pugh et al., 2000; Bitan et al., 2005, 2006; Heim et al., 2009). For example, Bitan et al. (2005, 2006) used a spelling task and a rhyming task in which participants determined whether the final word had either the same spelling or the same rhyme as either of the first two words, and found that the spelling task recruited stronger directional connections to the intra-parietal sulcus involved in orthographical form analysis and movement, and the rhyming task recruited stronger directional connections to left temporal cortex involved in phonological form analysis. To summarise, the extant literature has demonstrated that regional activation and inter-regional functional connection in visual word recognition are not purely stimulus-driven but highly task-guided instead.

The cognitive processes recruited for visual word recognition evolve over time rapidly (Serenó & Rayner, 2003; Carreiras, Armstrong, Perea, & Frost, 2014; Hauk, 2016). EEG and MEG signals with excellent temporal resolution are ideal tools for revealing the time course of visual word recognition. Some studies have shown that event-related potential differences between real words and pseudowords can be observed at between 150 ms and 200 ms after stimulus onset in a lexical decision task, indicating a very early lexicality effect (e.g., Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006; Hauk, Coutout, & Holden, & Chen, 2012; Araújo, Faisca, Barmao, Reis, & Petersson, 2015; Coch & Meade, 2016). These studies demonstrate the necessity of the application of electrophysiological measures in the examination of early neural dynamics (Carreiras et al., 2014). Electrophysiological evidence has also indicated that the influence of linguistic properties of visual words (e.g., word frequency, word length and imageability) on the time course of word recognition can be modulated by tasks with different cognitive goals within 300 ms from stimulus onset (e.g., Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Fujimaki et al., 2009; Strijkers, Yum, Grainger, & Holcomb, 2011; Grainger, Bertrand, & Grainger, 2015; Chen, Davis, Pulvermüller, & Hauk, 2015; Yum, Law, Su, Lau, & Mo, 2014; Mahé, Zesiger, & Laganaro 2015; Wang & Maurer, 2017). Strijkers et al. (2015) observed that when participants were required to decide whether a visual word was an animal word (semantic categorisation task), event-related potential differences between high-frequency and low-frequency words emerged from 120 ms onward; but when participants were required to decide whether the colour of a visual word was blue (colour categorisation task), the emergence of the differences was delayed by about 100 ms; that is, the differences started at 220 ms. The authors concluded that the speed and quality of lexical processing depends on the top-down intention: the specific goal of a task guides the time-course of lexical processing.

With excellent temporal resolution and good spatial resolution, MEG can also be used to reveal the evolution of inter-regional causal connections over time in the brain. Using a personal name detection semantic categorisation task, Woodhead et al. (2014) recently applied DCM to MEG data to investigate the causal influences from the frontal cortex to the left vOT within 300 ms after stimulus onset during visual word recognition. The results showed that, compared with meaningless symbols, visual words evoked a stronger feedback connection from the frontal cortex to the vOT in the first 200 ms. The finding strongly indicates that word recognition involves modulation from high-level frontal cortex on the low-level vOT even at the early stages. To further elaborate the nature of this early feedback, we recently conducted an MEG study in which we used different types of stimuli with different linguistic properties (Li et al., in preparation; Chapter 2). Specifically, we included real words, orthographically-legal pseudowords, consonant strings, and meaningless false fonts to see what types of linguistic properties trigger this early top-down feedback. Similar to the personal name detection task in Woodhead et al. (2014), the task we used was a semantic categorisation task in which participants decided whether a visual stimulus was an animal word. Consistent with Woodhead et al. (2014)'s main observation, we found a stronger feedback connection from the frontal cortex to vOT in the first 200 ms for real words than for false fonts. In addition, we found that the LIFG-to-LvOT feedback connection was stronger for real words than for pseudowords in both 1-150 and 1-200 ms time-windows, and that the feedback connection was stronger for pseudowords than for consonant strings in the 1-200 ms time-window. On the basis of these findings, it was concluded that both lexical-semantic and phonological information derived from visually presented words can drive early feedback from frontal cortex to vOT, and that lexical-semantic information begins to have an influence earlier than phonological information.

However, inter-regional effective connectivity in the brain appears to be modulated by specific task goals, as summarised in the fMRI studies above (e.g., Bitan et al., 2005, 2006; Heim et al., 2009). Therefore, it is likely that the pattern of inter-regional connectivity at the early stages of visual word recognition, as found in our previous study, could also be task-modulated. The existing findings that the task-modulation of brain activity at the early stages of word recognition (e.g., Bentin et al., 1999; Fujimaki et al., 2009; Strijkers et al., 2011, 2015; Chen et al., 2015; Yum et al., 2014; Mahé et al., 2015; Wang & Maurer, 2017) also suggest the possibility that task goals modulate the influences of linguistic properties on the inter-regional directional connections from frontal to vOT cortex at the early stages of visual word recognition. The present study sought to shed light on the issue of whether the early top-down feedback observed at the early stages of visual word recognition are dependent on task demands. In the current study, a visual discrimination task in which participants were required to decide whether a visual stimulus was a string of hashes (e.g., ###, #####) was compared to the semantic categorisation task used in our previous study. Comparisons of neural processing when one task draws on high-level linguistic properties while another restricts these properties and emphasises visual feature detection have also been used in previous fMRI studies (Yang et al., 2012; Mano et al., 2013; Yang & Zevin, 2014). However, these studies could not provide temporal information due to the poor temporal resolution of BOLD signals. Unlike these studies, the focus of the current study was on the early neural interactions between LIFG and LvOT.

To allow a direct comparison with the previous study (Li, et al., in preparation; Chapter 2), we kept the same the critical stimuli and the general experimental procedure including the apparatus, and matched age and reading ability of participants. However, we replaced the animal words used in the previous study with strings of hashes, and changed the task to that of visual discrimination. Specifically, in the visual discrimination task, participants responded by

pressing a button if the stimulus was a string of hashes; otherwise, participants did nothing. Thus, any difference between the two experiments would reflect the modulation by task. As in Li et al. (in preparation; Chapter 2), we specifically focused on the real words vs. pseudowords comparison, the pseudowords vs. consonant strings comparison, and the consonant strings vs. false fonts comparison. By doing so, we could clarify what type of effects – lexicality, phonological, and letter effects - specifically trigger early top-down feedback. Because no linguistic information is included in hashes, and the task could be performed by using only visual features, the need for high-level linguistic information would be expected to be minimised or even removed. By contrast, during the semantic categorisation task in which participants decided whether a stimulus was an animal word, high-level lexical-semantic processes must be drawn upon to make a decision. The comparison between these two tasks could answer how the influences of linguistic properties in early top-down feedback, as revealed in a linguistic task, are reflected in a non-linguistic task. In this study, we hypothesised that if one of the three pre-defined effects on the early top-down feedback from LIFG to LvOT is task-dependent, the feedback for that effect would be weakened or diminished in the visual discrimination task because it explicitly places emphases on visual features but not any linguistic properties of the stimuli and thus the connection between the LIFG and LvOT is not necessarily used to execute the task (see Mano et al., 2013). However, if an effect on the early top-down feedback from LIFG to LvOT is task-independent, we expected that the feedback for that effect would show some similarities in the two tasks.

Methods¹

Participants

Fifteen native English speakers (10 females, 5 males; age range: 19-31 years; mean age: 25 years) were recruited from Macquarie University campus to participate in this study. All participants reported that they had normal hearing, normal or corrected-to-normal vision and had no history of neurological disorders. The Edinburgh Handedness Inventory (Oldfield, 1971) was used to assess their handedness: fourteen were right-handed and one left-handed. The Sight Word Efficiency (SWE) and Phonemic Decoding Efficiency (PDE) subscales of the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner & Rashotte, 1999) were used to confirm that all participants had reading competence within the normal range (mean of overall standard score: 100 ± 13.10). All participants had normal performance IQ (mean of overall standard score: 113 ± 8.55) as measured by the Matrices subscale of the Kaufman Brief Intelligence Test, (Second Edition; KBIT-2; Kaufman & Kaufman, 2004). The current study was approved by the Human Research Ethics Committee (Medical Sciences) at Macquarie University. Prior to the beginning of the experiment, a signed consent form was obtained from every participant.

Overall Procedure

Participants completed the 80-minute MEG experiment first and then the 30-minute behavioural testing which included the reading ability test (TOWRE) and the intelligence test, (KBIT) either on the same day or the following day.

¹ In this study, the four critical conditions, experimental procedure, apparatus, and DCM data analysis were the same as the previous study (Li et al., in preparation; Chapter 2).

Materials

All critical stimuli were completely adopted from our previous MEG study (Li et al, in preparation; Chapter 2). The stimuli belonged to four categories: real words (RW), pseudowords (PW), consonant strings (CS) and false fonts (FF). Within each category there were 120 exemplars divided equally into four stimulus lengths of three, four, five and six characters. All stimuli were presented in Calibri in lower case and size 50. RWs were carefully chosen from MCWord (an Orthographic Wordform Database). PWs were firstly generated based on real words through the WUGGY software (Keuleers & Bysbaert, 2010. For details, see website <http://crr.ugent.be/programs-data/wuggy>) and then were matched to RWs in bigram frequency, trigram frequency and Coltheart's N based on MCWord database. See Table 1 in our previous MEG study for details (Chapter 2). CSs were also initially generated by the WUGGY software, then their bigram/trigram frequency and Coltheart's N were kept zero to make CSs as pure letter strings without any higher-level linguistic information. As in a previous study (Woodhead et al., 2014), FFs were direct translations of the real words using the "Carian" font (Jane Warren, personal communication). "Carian" fonts are characters adapted from the alphabet of an obsolete Anatolian language (Melchert, 2004). Some physical properties of the Carian script are similar to the modern Roman alphabetic characters, consisting of simple combinations of curved and/or linear elements. To reduce possible influences of similarity to English letters, some characters whose shape was visually similar to English letters, were replaced by others. Therefore, the FFs had no meanings and could not be pronounced.

MEG Experiment Procedure

The entire MEG experiment procedure included preparation, a practice run, the actual experiment and rests between blocks. The actual experiment consisted of six blocks. Each

block lasted about 6 minutes. Each block contained 80 critical trials (20 trials for each condition) and 16 catch trials - strings of hash symbols (e.g., ###, ##### and #####). Thus, the experiment included 480 trials (120 trials for each condition) and 96 strings of hash symbols. The order of stimuli was pseudo-randomised and presented on the screen in the magnetically shielded room to the participant using E-Prime 2.0 (<https://www.pstnet.com/eprime.cfm>). The distance between the screen and participants' eyes was 60 cm. On each trial, a black screen was initially presented with a white cross for visual fixation; then a stimulus was presented for 500 ms, followed by a white cross which was displayed for between 2800 ms and 3400 ms (see Figure 1). Participants were asked to attend to each trial carefully and respond with a button press when a string of hash symbols was presented on the screen. Catch trials were removed from the data analysis. Before the experiment, a practice run was conducted to familiarise participants with the task.

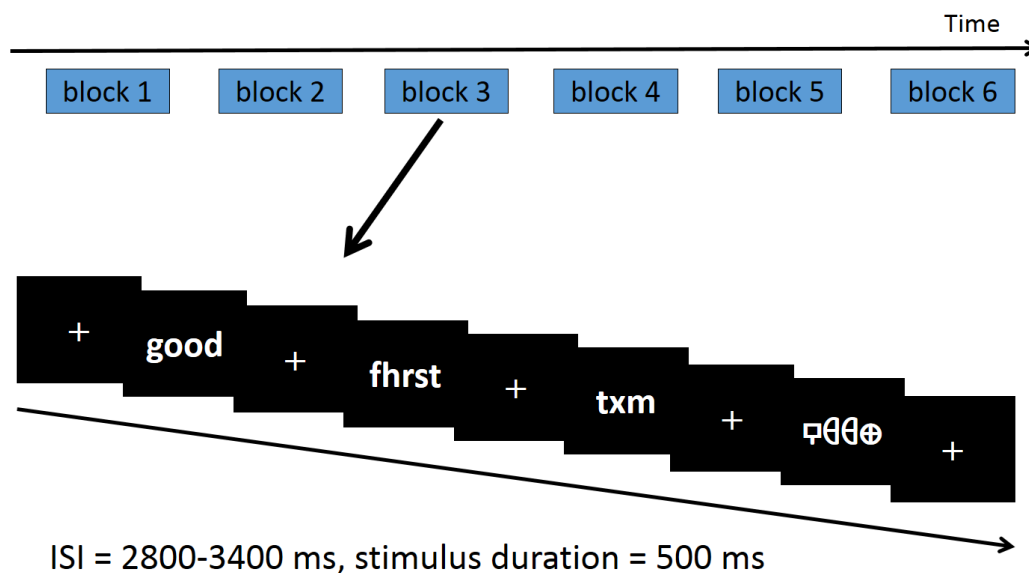


Figure 1. Experimental procedure for the study. Six blocks were included, with each having 96 trials including 80 critical trials and 16 hash strings.

MEG Data Acquisition and Pre-processing

MEG data were recorded using the KIT-Macquarie MEG160 (Model PQ1160R-N2, KIT, Kanazawa, Japan) located at Macquarie University, with participants lying in a magnetically shielded room. Data were recorded using 160 coaxial first-order gradiometers with a 50-mm baseline (Kado et al. 1999; Uehara et al. 2003). Prior to MEG data acquisition, the locations of three head landmarks (nasion and bilateral preauricular points), five marker coil positions, and the subject's head shape were recorded by 3D digitisation (Polhemus Fastrack, Colchester, VT). Participants' head position and corresponding sensor positions were measured by energising the five marker coils briefly. The head motion was less than 6 mm for each participant in each block during the entire recording. The online sample rate was 1000 Hz.

MEG data analysis was conducted using SPM12 software (Litvak et al., 2011) implemented in Matlab 2014b (The MathWorks, Inc.). The analysis steps were as follows. The original data sampled at 1000 Hz were high-pass filtered at 1 Hz and then low-pass filtered at 30 Hz. The filtered data were epoched from 150 ms before stimuli to 600 ms after stimuli. The pre-stimulus time window (-100 - 0 ms) was used for baseline correction. To reject artefacts, the Fieldtrip visual artefact rejection toolbox, which expresses every time point as a deviation from the mean over all time and channels, was used to remove extreme trials in four conditions for each participant based on the variance within each channel (http://www.fieldtriptoolbox.org/reference/ft_rejectvisual). No more than 5% of trials in any condition were excluded by the artefact rejection procedure. Finally, robust averaging was applied to the epoched data across trials within each condition (RW, PW, CS and FF) (Litvak et al., 2011; Wager, Keller, Lacey, & Jonides, 2005).

A head model for constructing source activity in the DCM analysis for each participant was built based on several settings below. First, a normal-resolution cortical mesh generated

from the MNI template (ICBM512) was created. Then, three MEG fiducial labels (nasion, left pre-auricular point (LPA) and right pre-auricular point (RPA)) and headshape points were used to coregister the MEG data. Finally, a single shell was used to define the forward model in the following DCM analysis.

Dynamic Causal Modeling (DCM)

DCM is a widely-used and powerful tool for examining the causal relationships between brain areas (effective connectivity; the influence a region exerts on another region). DCM allows investigation of the effects generated by stimulus types and/or how task demands modulate the strength of the causal connections between brain areas. The details and basic principles of the DCM for ERP data have been extensively described elsewhere (David et al., 2006; Kiebel, Carrido, Moran, Chen, & Friston, 2009).

Basic steps of DCM. A standard pipeline of DCM analysis includes the following parameters that need to be specified: the definition of the source locations; the definition of the connections between these sources; the source(s) of input(s); the connections to be modulated by variations in stimulus types and/or tasks. A DCM model can be established through changing these parameters iteratively until the predicted neural activity produced by the model best matches the observed data. That also means that one can use this approach to estimate a great body of models by changing the number of sources, the connections between sources and sensory inputs, to test a specific hypothesis. Bayesian statistics are performed to investigate which model provides the best explanation of the data (Bayesian model selection, BMS; Penny, Stephan, Mechelli, & Friston, 2004). If there is no model that is apparently the best one, then a weighted average model can be established via the approach of Bayesian Model Averaging (BMA; Penny et al., 2010). As used in the previous chapter, here we used a BMA approach to estimate connection strength.

Nodes of DCM. Three left-hemisphere regions (left middle occipital cortex (OCC), left ventral occipito-temporal cortex (vOT) and left inferior frontal gyrus (IFG) were obtained from an ALE meta-analysis based on eight MEG/EEG studies of visual word recognition (see Li et al., in preparation; Chapter 2). The vOT coordinates were very close to those reported by previous studies (e.g., Cohen, Jobert, Le Bihan, & Dehaene, 2004; Dehaene & Cohen, 2011). The right-hemisphere homologues of these areas were also obtained based on these coordinates. Thus, a six-node network for the DCM analysis was constructed which included bilateral OCC (MNI, ± 26 -94 6), bilateral vOT (MNI, ± 46 -56 -18) and bilateral IFG (MNI, ± 58 8 22). Studies with similar experimental settings have shown that this six-node network provides the best fit to the MEG data (Woodhead et al., 2014), thus its nodes were considered as the sources of the DCM spatial model in the following BMA analysis.

DCM models. The total number of possible connections in a model with 6 nodes is 30, and the total number of possible combinations of those connections (the total model space) is 2^{30} . To reduce computational load, some constraints had to be enacted to limit the size of the model space. Following the rules of generating DCM models in the previous experiment, the number of independent connections was reduced to 5 in the model, creating a total model space of 32 models for each comparison for each participant; each modelling a different combination of connections between sources mediating trial-specific effects.

Following the previous experiment (see Chapter 2), to better examine the changes in connection strengths over time, the DCM analysis was conducted in three different time-windows, 1-100 ms, 1-150 ms and 1–200 ms (see Garrido, Kilner, Kiebel, & Friston, 2007, Woodhead et al., 2014 and Poch et al., 2015 for a similar approach). Three peaks at around 100 ms, 150 ms and 200 ms are also very clear in the early time-windows of the ERF. These three time-windows were chosen based on previous studies in which the main reading areas in the left hemisphere were clearly activated in the first 200 ms during visual word recognition (e.g.,

Pammer et al., 2004; Cornelissen et al., 2009; Wheat, Cornelissen, Frost, & Hansen, 2010; Woodhead et al., 2014).

Bayesian Model Averaging (BMA). All 32 models were estimated in three time-windows for each participant. The 32 models were averaged by group-level BMA with random effects (Penny et al., 2010) to investigate the average strength of the stimulus-modulated connections, i.e., the gains for the four comparisons: RW vs. PW (lexicality effect), PW vs. CS (phonological effect), CS vs. FF (letter effect) and RW vs. FF. For each connection, an average gain value equal to 0 indicates that no stimulus-type effect occurs in the connection; an average gain value significantly greater than 0 indicates that the connection strength is stronger for the former stimulus type; and vice versa, an average gain value significantly smaller than 0 indicates that the connection strength is stronger for the latter stimulus type.

To test whether the gain values from the BMA analysis were significant or not, a non-parametric proportion test was adopted. For each connection, the distribution of the gain was reconstructed by generating 100,000 samples from a Gaussian distribution based on the posterior mean and standard deviation calculated in the BMA step. The rule was that if more than 90 percent of the samples were greater or smaller than zero, the connection was considered to be significantly stronger for the former condition or the latter condition, respectively; otherwise, the connection strength was considered to be equal between two conditions in the connection. A similar approach has been used before (Seghier, Josse, Leff, & Price, 2011; Richardson, Seghier, Leff, Thomas, & Price, 2011; Woodhead et al., 2014; Xu, Wang, Chen, Fox, & Tan, 2016; also see our previous study).

Direct comparisons between the tasks. To estimate task-modulation effects of the key early feedback connection, i.e. the LIFG-to-LvOT connection, the connection strength was compared between two tasks (semantic categorisation task and visual discrimination tasks) for each comparison (RW vs. FF, RW vs. PW, PW vs. CS and CS vs. FF). First, 100,000 samples

from a Gaussian distribution based on posterior mean and standard deviation were generated for each connection in each comparison. Then the differences between the two tasks for each comparison was assessed by the fraction of samples that are higher in the semantic categorisation task than the visual discrimination task or opposite direction. If an effect was reported at a posterior probability threshold of 0.90, then the effect was considered to be significant (see a similar approach in Seghier et al., 2010).

Results

The behavioural results for task performance, the results of event-related fields at the sensor space, and the results of the DCM analysis for each comparison are reported below.

Task Performance

The averaged reaction time for catch trials, i.e. hash strings, was 490 ms ($SD = 69$ ms, range 392-594 ms). The average accuracy for catch trials was 0.98 ($SD = 0.043$, range 0.83-1).

Event-related Fields (ERFs)

As used in the previous chapter, the root mean square (RMS) was used to depict the time course and strength of event-related fields at the sensor space. The averaged RMS values across four conditions were calculated in the first 500 ms after stimulus presentation. In Figure 2, event-related peaks at around 100 ms, 150 ms and 200 ms appeared in order as a function of time-course. The observation is in line with previous studies of visual word recognition (e.g., Hauk et al., 2006; Chen et al., 2013).

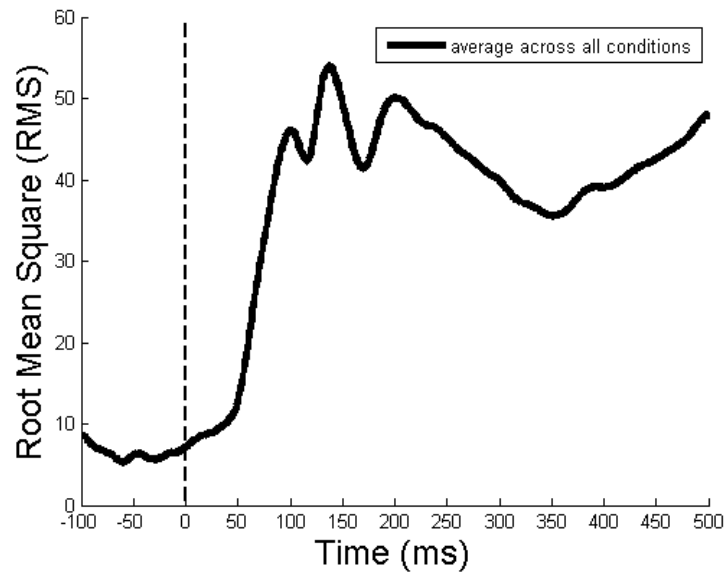


Figure 2. The average root mean square (RMS) plot across all conditions and all subjects as a function of time course during the first 500 ms after stimulus onset. Three event-related field peaks can be seen at 100 ms, 150 ms and 200 ms, i.e., M100, M150 and M200 components.

Dynamic Causal Modeling

The results of BMA analysis assessing the stimulus-type effects on the causal relationships between reading nodes in each time window are displayed in Figures 3, 4, 5 and 6 for each comparison. Table 1 provides the details of posterior means of each connection in each time-window. To examine the inter-regional connections in the reading network, our focus was on the connections among bilateral IFG, vOT and OCC. Self-connections and horizontal lateral connections are not provided here (see supplementary Table 1 for the details). We have defined RW vs. PW as a lexicality effect, PW vs. CS as a phonological effect and CS vs. FF as a low-level letter effect. The DCM results for each comparison were reported below.

Table 2.

Posterior means for feedforward and feedback connections in different comparisons. Values in bold were significantly stronger for the former (> 0) or weaker for the latter (< 0) than the threshold (90 %) (Corresponding p values were smaller than 0.05).

	RW vs. FF (overall)			RW vs. PW (lexicity effect)		
	1-100 ms	1-150 ms	1-200 ms	1-100 ms	1-150 ms	1-200 ms
Feedforward						
LvOT to LIFG	-0.014	0.002	-0.019	0.000	-0.036	0.069
LOCC to LvOT	-0.086	0.200	0.100	-0.129	-0.083	-0.118
LOCC to LIFG	-0.096	0.113	0.035	-0.185	-0.033	-0.038
RvOT to RIFG	0.058	-0.027	-0.058	-0.017	-0.052	-0.003
ROCC to RvOT	0.003	-0.076	-0.131	-0.066	-0.131	-0.069
ROCC to RIFG	0.001	-0.155	-0.075	-0.083	0.018	-0.066
Feedback						
LIFG to LvOT	0.042	0.001	0.006	0.019	0.079	0.062
LvOT to LOCC	0.060	0.018	-0.047	0.018	0.046	-0.125
LIFG to LOCC	-0.013	0.001	-0.034	-0.020	0.032	-0.033
RIFG to RvOT	-0.041	0.048	0.137	0.054	-0.037	0.047
RvOT to ROCC	-0.014	0.032	-0.029	0.025	0.081	-0.054
RIFG to ROCC	0.008	-0.046	0.026	-0.026	0.053	-0.008

	PW vs. CS (phonological effect)			CS vs. FF (letter effect)		
	1-100 ms	1-150 ms	1-200 ms	1-100 ms	1-150 ms	1-200 ms
Feedforward						
LvOT to LIFG	-0.019	-0.037	0.024	-0.047	0.031	-0.003
LOCC to LvOT	0.045	-0.116	0.028	-0.057	0.199	0.148
LOCC to LIFG	0.188	-0.093	-0.053	-0.148	-0.086	0.059
RvOT to RIFG	-0.022	0.100	-0.004	0.009	0.017	-0.035
ROCC to RvOT	-0.135	-0.026	-0.206	0.066	0.116	-0.050
ROCC to RIFG	0.156	-0.064	-0.124	-0.155	-0.012	-0.067
Feedback						
LIFG to LvOT	-0.065	-0.028	-0.070	0.004	0.103	0.057
LvOT to LOCC	-0.071	-0.073	0.024	0.081	0.029	0.033
LIFG to LOCC	-0.035	-0.047	-0.047	0.013	0.046	-0.040
RIFG to RvOT	-0.056	-0.033	0.001	-0.018	0.004	-0.048
RvOT to ROCC	-0.023	-0.033	0.071	0.037	-0.120	0.042
RIFG to ROCC	0.037	-0.063	0.014	-0.004	-0.032	-0.093

RW vs. FF comparison (Overall). The main results of the RW vs. FF comparison are as follows (also see Figure 3 and Table 1 for details). (1) Only one feedback connection, i.e. the RIFG-to-RvOT connection was stronger for RW than for FF in the 1-200 ms time-window. (2) The LOCC-to-LvOT connection was stronger for FF than for RW in the 1-100 ms time-window, but was stronger for RW than for FF in both 1-150 ms and 1-200 ms time-windows. (3) The ROCC-to-RvOT and ROCC-to-RIFG feedforward connections were stronger for FF than for RW in the 1-150 ms and 1-200 ms time-windows.

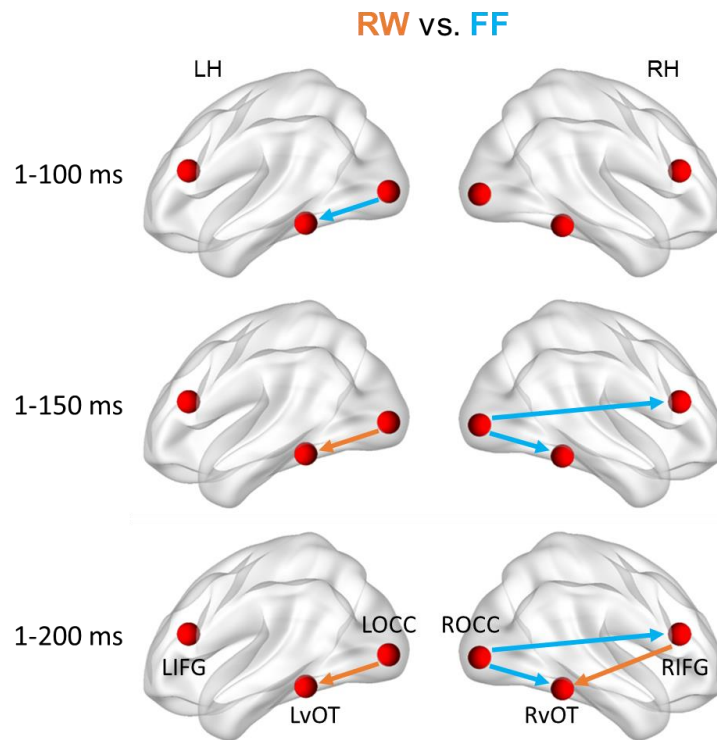


Figure 3. RW vs. FF comparison for inter-regional causal connections in three time-windows, 1-100 ms, 1-150 ms, and 1-200 ms. Only significant causal connections are shown here. The arrows denote directions of causal influences. An orange line denotes a connection that is stronger for RW than for FF while a blue denotes a connection that is stronger for FF than for RW.

RW vs. PW comparison (Lexicality effect). The main results of the RW vs. PW comparison are as follows (also see Figure 4 and Table 1 for details). (1) The LIFG-to-LvOT

feedback connection was stronger for RW than for PW in both 1-150 ms and 1-200 ms time-windows. (2) All other connections including the ROCC-to-RvOT and ROCC-to-RIFG feedforward connections in all three time-windows, the LOCC-to-LvOT and LOCC-to-LIFG feedforward connections in all three time-windows, and the LvOT-to-LOCC feedback connection in the 1-200 ms time-window were stronger for PW than for RW. (3) No significant feedback connection was found in the 1-100 ms time-window in the left hemisphere (LH), and all significant connections in three time-windows in the right hemisphere (RH) were feedforward connections.

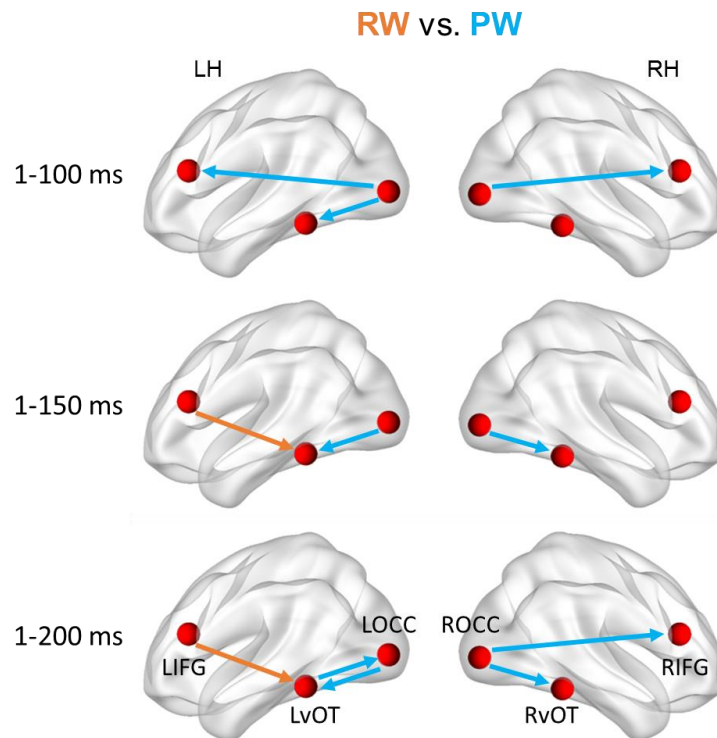


Figure 4. RW vs. PW comparison (lexicity effect) for inter-regional causal connections in three time-windows, 1-100 ms, 1-150 ms, and 1-200 ms. Only significant causal connections are shown here. The arrows denote directions of causal influences. An orange line denotes a connection that is stronger for RW than for PW while a blue denotes a connection that is stronger for PW than for RW.

PW vs. CS comparison (Phonological effect). The main results of the PW vs. CS comparison are as follows (also see Figure 5 and Table 1 for details). (1) Both the LIFG-to-LvOT feedback and the LOCC-to-LIFG feedforward connections were stronger for CS than for PW in the 1-200 ms time-window, which was followed by both the LOCC-to-LIFG and LOCC-to-LvOT feedforward connections stronger for CS than for PW in the 1-150 ms time-window. (2) No significant feedback connections were found in either the 1-100 ms or 1-150 ms time-windows for both hemispheres.

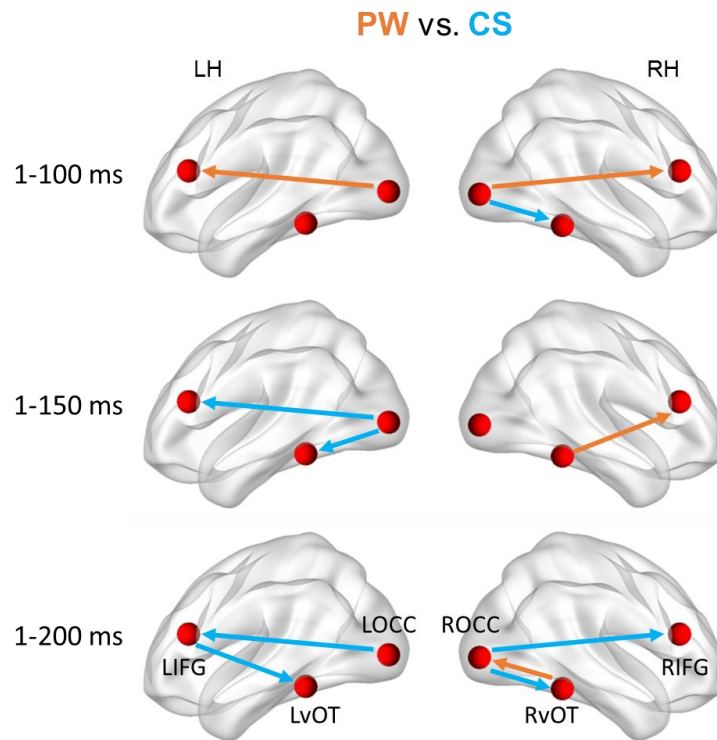


Figure 5. PW vs. CS comparison (phonological effect) for inter-regional causal connections in three time-windows, 1-100 ms, 1-150 ms, and 1-200 ms. Only significant causal connections are shown here. The arrows denote directions of causal influences. An orange line denotes a connection that is stronger for PW than for CS while a blue line denotes a connection that is stronger for CS than for PW.

CS vs. FF comparison (Letter effect). The main results of the CS vs. FF comparison are as follows (also see Figure 6 and Table 1 for details). (1) The LIFG-to-LvOT feedback

connection was stronger for CS than for FF in the 1-150 ms time-window, which was accompanied by the LOCC-to-LvOT feedforward connection being stronger for CS than for FF in the same window. (2) The bi-directional connections between RIFG and ROCC were stronger for FF than for CS. (3) Two feedforward connections, i.e. the OCC-to-IFG connections in both hemispheres were stronger for FF than for CS.

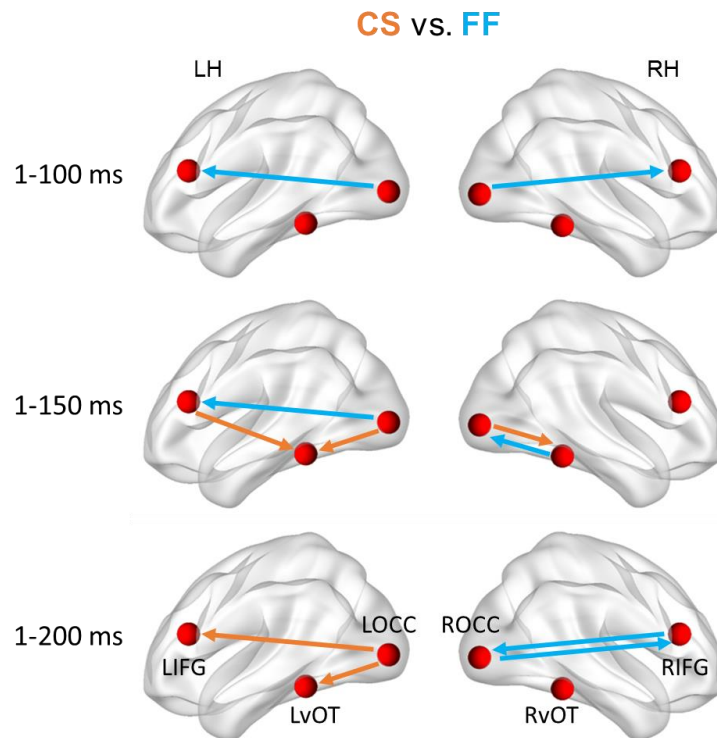


Figure 6. CS vs. FF comparison (letter effect) for inter-regional causal connections in the three time-windows, 1-100 ms, 1-150 ms, and 1-200 ms. Only significant causal connections are shown here. The arrows denote directions of causal influences. An orange line denotes a connection that is stronger for CS than for FF while a blue line denotes a connection that is stronger for FF than for CS.

Comparisons between two tasks. The results from the present experiment were compared with those of Chapter 2, where the semantic categorisation task was used. The average RTs of catch trials (animal words and hash strings) with correct response in the semantic categorisation task and the visual discrimination task were 659 ms ($SD = 89$ ms) and

490 ms ($SD = 69$ ms) respectively. The average accuracy of catch trials in the two tasks were 0.94 ($SD = 0.04$) and 0.98 ($SD = 0.043$) respectively. Two-sample t-tests (two-tailed) revealed that the mean RT for animal words in the semantic task was significantly longer than that of hash strings in the discrimination task ($t(28) = 5.80, p < .001$), and the accuracy of animal words in the semantic task was significant lower than that of hash strings in the discrimination task ($t(28) = -2.39, p = .02$). The average RMS values reflecting event-related field strength in the semantic categorisation and visual discrimination tasks were compared with a two-sample t-test (two-tailed). The results showed that the event-related field strength from 50 ms to 77 ms after stimulus onset was stronger in the discrimination task than in the semantic task (see Figure 7).

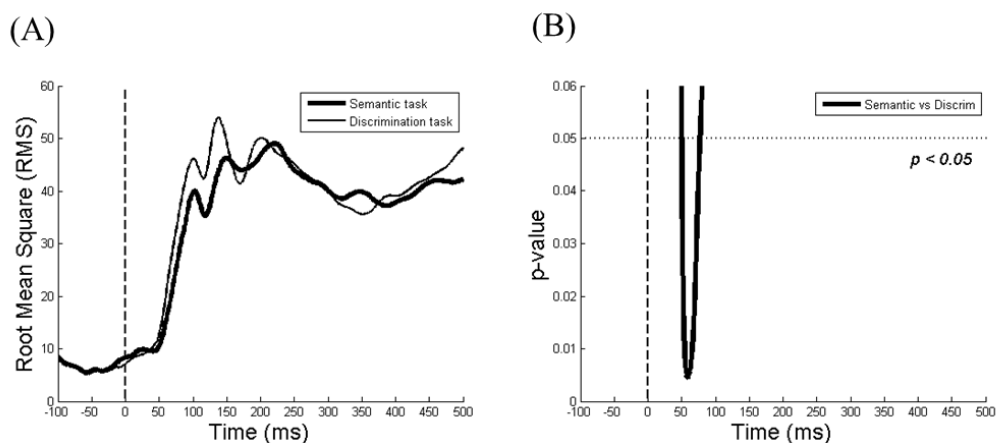


Figure 7. (A) The average root mean square (RMS) plot across all conditions and all subjects as a function of time course during the first 500 ms after stimulus onset in the semantic categorisation task (the thick curve) and the visual discrimination task (the thin curve). *(B)* P-values as a function of time evolution obtained with a two-sample t-test (two-tailed) on the RMS values between two tasks. The dotted line denotes the boundary of the 0.05 p-value.

The Bayesian Model Averaging (BMA) results from the two tasks were put together to examine effects of task modulation on the early LIFG-to-LvOT feedback connection (see Figure 8). Statistical analyses for each comparison showed that (1) the differences between

tasks in RW vs. FF in each time-window were not significant ($p > 0.05$); (2) the differences between tasks in RW vs. PW in each time-window were not significant ($p > 0.05$); (3) the difference between tasks in PW vs. CS in the 1-200 time-window was significant ($p < 0.01$); (4) the differences between tasks in CS vs. FF in the 1-150 and 1-200 ms time-windows were significant ($p < 0.01$). These results showed that the lexicality effect defined by RW vs. PW, in the early LIFG-to-LvOT feedback connection was independent of task; however, the overall effect defined by RW vs. FF, the phonological effect defined by RW vs. CS, and the letter effect defined by CS vs. FF, in the early LIFG-to-LvOT feedback connection, are all dependent on task.

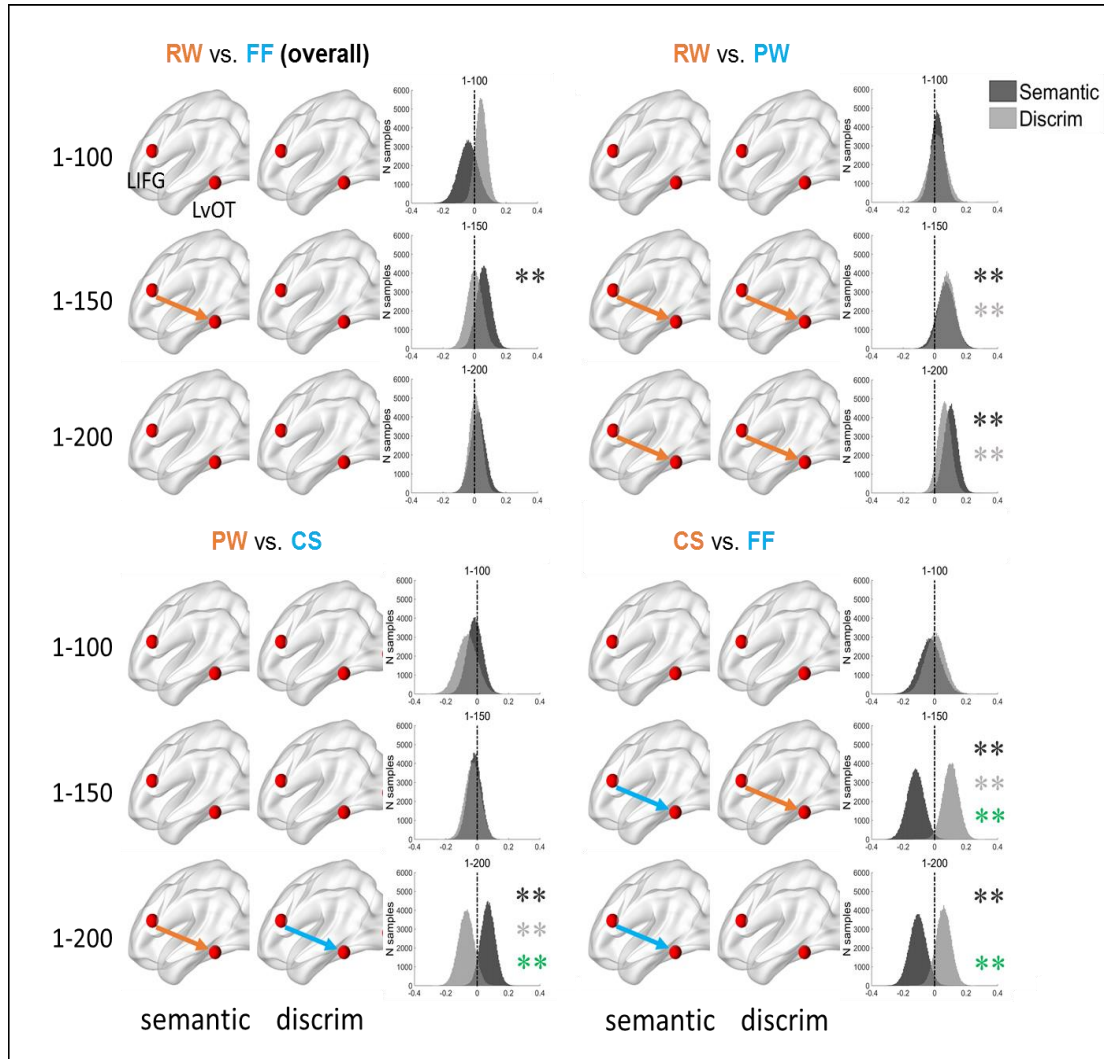


Figure 8. Effects of the four different comparisons for the LIFG-to-LvOT connection in three time-windows in the semantic categorisation and visual discrimination tasks, and the differences between the tasks. **Left side** in each comparison, strength of the LIFG-to-LvOT connection in three windows. An orange line denotes the connection that is stronger for the former condition than for the latter condition while a blue line denotes the connection that is stronger for the latter condition than for the former condition. **Right side** in each comparison, distributions of 100,000 samples of the LIFG-to-LvOT connection generated by posterior means and corresponding standard deviations for each window in each task. Dark grey colour denotes the semantic categorisation task and light grey colour denotes the visual discrimination task. Dashed lines denote zero value. Dark grey asterisks denote significant effects in the semantic categorisation task; light grey asterisks denote significant effects in the visual discrimination task; green asterisks denote significant differences between the two tasks. ** $p < 0.01$.

Discussion

Following the previous study, this study sought to examine task modulation of the early LIFG-to-LvOT top-down feedback during visual word recognition. This was done by carrying out the same basic experiment as reported in Chapter 2 (Li et al., in preparation) but with a visual discrimination task rather than a semantic categorisation task, with effects of lexicality, phonological processing, and letter processing being the focus. It was expected that if an effect on the early top-down feedback is task-dependent, the feedback for that effect would be weakened or diminished in the discrimination task. We found that the results in relation to task modulation were dependent on which comparison was examined. The LIFG-to-LvOT feedback connection was stronger for real words than for pseudowords in both 1-150 ms and 1-200 ms time-windows, indicating a lexicality effect; and this connection was also stronger for consonant strings than for pseudowords in the 1-200 ms time-window and for false fonts in the 1-150 ms time-window, indicating a phonological effect and a letter effect, respectively. Comparing the results across the two experiments based on the posterior means of the LIFG-to-LvOT connection revealed that the basic pattern of the real words vs. pseudowords comparison was almost same in the two tasks, indicating that the lexicality effect is task-independent. However, the patterns of the pseudowords vs. consonant strings and consonant strings vs. false fonts comparisons in the discrimination task were almost the opposite to those in the semantic task, indicating that the phonological and letter effects are task-dependent. In addition, the basic finding that the top-down feedback was stronger for real words than for false fonts in the semantic task - replicating Woodhead et al (2014) - was absent in the discrimination task. These results suggest that the influences of lexicality on the early LIFG-to-LvOT top-down feedback are likely task-independent or automatic, whereas the influences of phonological and letter effects are highly task-dependent or flexible. These observations are discussed in detail below.

Task Dependence of Phonological and Letter Effects on the Top-down Feedback

It was interesting to observe that the phonological and letter effects on the LIFG-to-LvOT connection in the present discrimination task and the previous semantic task (Li et al., in preparation; Chapter 2) showed almost opposite patterns. Considering the fact that the critical stimuli and experimental settings were kept the same in the two experiments, the reason why the patterns of the LIFG-to-LvOT top-down feedback in the two tasks were almost opposite is likely attributable to the nature of the task the participants performed. Specifically, the differences in this feedback connection between tasks may originate from the distinct cognitive demands and aims of the two tasks. Studies examining the task modulation of cognitive and neural correlates of visual word processing have previously demonstrated that a specific task magnifies relevant cognitive and neural resources to serve its aims (e.g., Balota et al., 2004; Bitan et al., 2005, 2006; Yang et al., 2012; Mano et al., 2013; Strijkers et al., 2015; McNorgan et al., 2015). In the current case, the semantic task puts high emphasis on high-level lexical-semantic access, whereas the discrimination task is simply to decide whether a string is a hash string and thus it relies more on the detection of very low-level visual features. The finding that the early even-related field strength was stronger for the discrimination task than the semantic task (see Figure 7) demonstrates that the discrimination task indeed relies more on low-level visual feature detection.

Due to the high emphasis on lexical-semantic access in the semantic task, participants had to construct phonological information from pseudowords they have never seen before and then used phonological information to confirm that these pseudowords similar to real words do not have meanings and thus were not animal words. However, it is relatively easier to reject consonant strings as being animal words. Thus, it is likely that the difficulty of rejecting pseudowords as being animal words led to a stronger LIFG-to-LvOT connection for

pseudowords than for consonant strings in the semantic task. Here, the cognitive emphasis was biased toward pseudowords rather than consonant strings. In contrast, cognitive resources were directed to the processing of low-level visual features when the discrimination task was performed. The LIFG is an area for phonological and lexical-semantic processing (Taylor et al., 2013; Carreiras et al., 2014) and is recruited to differentiate pseudowords from consonant strings, so this is consistent with the observation that the LIFG-to-LvOT connection was stronger for consonant strings than for pseudowords in the discrimination task. Here, the cognitive emphasis was biased to consonant strings rather than pseudowords.

Past studies have also found that the LIFG is activated more by letters than by symbols or colours, indicating its involvement in letter recognition (e.g., Flowers et al., 2004; Liu et al., 2010, 2011; Vartiainen, Liljestrom, Koskinen, Renvall, & Salmelin, 2011). Liu et al. (2011) further found top-down feedback from the LIFG to visual cortex in letter recognition. Here we also observed this effect in the top-down LIFG-to-LvOT feedback for letter strings in the discrimination task. Because consonant strings and false fonts are all meaningless visual strings and false fonts are visually similar to hash strings, participants had to confirm the differences in the basic visual features between hash strings and false fonts before they rejected consonant strings as being hash strings. This led to the observation that the LIFG-to-LvOT connection was stronger for consonant strings than for false fonts in the discrimination task. However, the feedback connection showed an opposite pattern in the semantic task: it was stronger for false fonts than for consonant strings in both the 1-150 ms and 1-200 ms windows. At the same time, the RIFG-to-RvOT connection was stronger for consonant strings than for false fonts. A possible explanation is that consonant strings recruited more right-hemisphere neural resources flowing from frontal to vOT to process surface properties of letter combinations when the semantic task was performed. Due to the demand for high-level linguistic processing in the semantic task, the bilateral IFG were likely jointly involved in top-down modulation to the

vOT of each respective hemisphere and these two top-down routes are complementary to each other. However, there was no need for the recruitment of right-hemisphere top-down feedback in the discrimination task because of the simplicity of the discrimination between consonant strings and false fonts.

In addition to the task modulation of the phonological and letter effects, the overall effect of the real words vs. false fonts comparison on the LIFG-to-LvOT connection was also task-dependent. The stronger LIFG-to-LvOT connection for real words than for false fonts in the 1-150 ms time-window in the semantic task disappeared in all three time-windows in the discrimination task. Both Woodhead et al. (2014) and Li et al. (in preparation, Chapter 2) found that this early feedback connection was stronger for real words than for false fonts. Both studies used a semantic categorisation task; Woodhead et al. used a personal name detection task and Li et al. used an animal word detection task. Therefore, it is likely that the high lexical-semantic demands of these two tasks resulted in a similar observation on the LIFG-to-LvOT top-down feedback. However, the stronger feedback for words than for false fonts was washed out by the discrimination task which had non-linguistic cognitive demands. This result strongly indicates that the nature of the task - linguistic or non-linguistic - contributes to the early top-down influences from inferior frontal to ventral occipito-temporal cortex. Top-down feedback is not automatic and free from task goals, but instead is highly modulated by them.

Note that the absence of this significant difference between the overall real words and false fonts comparison in the discrimination task reflects that the top-down feedback in the two conditions is comparable. Real words also have lexical information, but false fonts do not. This seemingly is in contradiction with the observation that lexical influence on the top-down feedback was present in the discrimination task because the feedback was stronger for real words than for pseudowords (see the section below). A detailed inspection of the nature of the discrimination task tends to support that they do not contradict each other. Due to the similarity

in surface properties between false fonts and hash strings and the reliance of the discrimination task on surface properties, the top-down feedback had to be heavily biased to false fonts so that a correct decision – reject false fonts as being non-hashes – could be made. It is likely that the top-down feedback was strong for both real words and false fonts, but the differences between them were removed by a strong bias toward false fonts in the discrimination task because of its reliance on surface properties. Therefore, this finding does not contradict the task-independence of lexical influence on the feedback *per se*.

Task-Independent Lexical Influences on Top-down Feedback

Although the phonological and letter effects in the semantic categorisation and visual discrimination tasks showed opposing patterns, the lexicity effect on the early top-down feedback connection was very consistent across the two tasks (see Figure 8). Specifically, the early lexicity effect on the feedback connection also appeared in both the 1-150 ms and 1-200 ms time-windows when the discrimination task was performed, and its magnitude was almost equal to that in the semantic task. Considering the earliness and consistency of the lexicity effect in both tasks, the role of lexicity in the early feedback from LIFG to LvOT is likely to be highly automatic and thus be task-independent. In the extant literature, an early lexicity effect at around or even before 150 ms from stimulus onset has been previously revealed by electrophysiological measures (Sereno, Rayner, & Posner, 1998; Carreiras et al., 2014). Specifically, event-related responses to real words and pseudowords start to diverge in the classic N170 time-window (e.g., Sereno et al., 1998; Hauk et al., 2006, 2012; Mahé et al., 2012; Araújo et al., 2015; Coch & Meade, 2016). For example, Hauk et al. (2006) found that differences in event-related responses appeared at 160 ms after word onset when a lexical decision task was performed. The current study extends findings of the earliness of lexicity effects on the time course of visual word recognition observed in previous work (Hauk et al.,

2006). The M/EEG studies mentioned above generally used high-level linguistic tasks, such as lexical decision and semantic categorisation. In terms of linguistic demands, these tasks, to some extent, are similar to the semantic categorisation task used in Li et al. (in preparation, Chapter 2) because all of them emphasise lexical-semantic access. However, these tasks are very different from the visual discrimination task, but it was found that the lexicality effect also occurred in this task.

Why was the lexicality effect on the LIFG-to-LvOT top-down feedback free from the influences of task goals? A crucial factor, reading proficiency, could contribute to the task independence of this lexicality effect. Both studies recruited young adult readers who were university students with more than 10 years reading experience and thus could be considered expert readers. The direct connections between orthography and semantics can be very strong and automatic in these readers such that when a word is displayed, the meaning can be rapidly extracted (e.g., Coltheart et al., 2001) and may exert early top-down feedback no matter what tasks are performed. Furthermore, real words are highly familiar to these readers and their memory traces have been well-established; but pseudowords are totally new and no memory traces in the brain have been established. Thus, the existing memory traces for real words can be rapidly activated for top-down feedback, whereas these are almost absent for pseudowords. Interestingly, with a lexical decision task and a symbol detection task in fMRI studies, Yang et al. (2012) and Yang and Zevin (2014) found that the activation in the LIFG evoked by real characters was comparable in the two tasks. Yang et al. (2012) further observed this pattern in the LvOT. These studies suggest that the activation in the LIFG is highly automatic at least in a task with no explicit linguistic demands, which is consistent with the present finding of task-independent lexicality effects in top-down feedback.

Overall Connectivity Pattern

Similar to the findings in the semantic task, significant left-lateralised top-down feedback was also found in the real words vs. pseudowords and pseudowords vs. consonant strings comparisons in the discrimination task, indicating that the left lateralisation of the word recognition brain network is generally independent of task. Past fMRI studies have reported a similar observation in both linguistic tasks (e.g., Taylor et al., 2013) and non-linguistic tasks (e.g., Vinckier et al., 2007; Liu et al., 2008). Both Vinckier et al. (2007) and Liu et al. (2008) found that a left-lateralised word-recognition network in a non-linguistic task. Furthermore, our previous MEG study using the semantic task found that the predominance of the LIFG-to-LvOT feedback connection in the 1-150 ms and 1-200 ms time-windows and the predominance of the OCC-to-vOT and vOT-to-IFG feedforward connections in the 1-100 ms time-window (Li et al., in preparation; Chapter 2), which is consistent with the pattern in the discrimination task. These consistent findings indicate that both tasks shared a similarity in that significant feedforward connections occurred earlier than the significant LIFG-to-LvOT feedback connections. This suggests that relevant information on visual words is firstly processed in a purely bottom-up hierarchical manner (information flows from low-level to high-level areas), and this is then followed by joint bottom-up and top-down processing; that is, information flows from low-level to high-level areas coupled by information flowing from high-level to low-level areas (also see Discussion in Chapter 2). These similarities between the two tasks indicate that tasks with different emphases but with the same stimulus inputs still preserve the general neural hierarchy of word recognition.

Limitations

The current DCM study comes with some limitations. One is that we compared the DCM results in two different tasks performed by two separate but well-matched samples. The reason why we adopted this approach is that fatigue could affect the neural responses to stimuli if participants were required to finish two tasks with each lasting about 80 minutes. However, this between-subject design still may contribute as a confounding factor because of small sample size in each task and potential large inter-subject variability. Previous fMRI studies of reading with two different tasks have adopted a within-subject design; that is, all participants finished two tasks (Yang et al., 2012; Mano et al., 2013; Chen et al., 2015; Yang & Zevin et al., 2014; Strijkers et al., 2015). Future studies could use a within-subject design to test the reliability and robustness of task modulation on early top-down feedback revealed in this study.

Conclusion

Following the previous study, this study applied DCM analyses to MEG data to examine the early feedback from frontal to ventral occipito-temporal cortex during visual word recognition in a visual discrimination task with no linguistic demands, and then compared the results with those in a semantic categorisation task. It was found that the LIFG-to-LvOT top-down feedback was highly dependent on the effects that were examined. The earliness and consistency of the lexicality effect in the two tasks indicates the automatic influences of lexical-semantic factors on the LIFG-to-LvOT top-down feedback during visual word recognition. The phonological and letter effects on the top-down feedback in the two tasks were different from each other, providing strong evidence for the task-dependence of these effects. The consistent division of labour for feedback and feedforward connections, and similar lateralisation of word recognition in the brain in the two tasks, further suggest that tasks with different cognitive aims

do not change the general neural hierarchy. The findings observed here have implications on the roles of task goals in modulating inter-regional information flow in the brain during visual word recognition.

Appendix A

Supplementary Table 1.

Posterior means for feedforward, feedback, lateral, and self-connections in different comparisons. Values in bold were significantly stronger for the formers (> 0) or weaker for the latters (< 0) than the chance (90 %) (Corresponding p values were smaller than 0.05).

Feedback	RW vs. FF					
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LIFG to LvOT	0.042	0.036	0.001	0.048	0.006	0.039
LvOT to LOCC	0.060	0.084	0.018	0.071	-0.047	0.064
LIFG to LOCC	-0.013	0.064	0.001	0.047	-0.034	0.053
RIFG to RvOT	-0.041	0.034	0.048	0.048	0.137	0.039
RvOT to ROCC	-0.014	0.085	0.032	0.075	-0.029	0.064
RIFG to ROCC	0.008	0.064	-0.046	0.047	0.026	0.053
<i>Feedforward</i>						
LvOT to LIFG	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LOCC to LvOT	-0.014	0.052	0.002	0.052	-0.019	0.058
LOCC to LIFG	-0.086	0.062	0.200	0.049	0.100	0.046
RvOT to RIFG	-0.096	0.044	0.113	0.041	0.035	0.044
ROCC to RvOT	0.058	0.050	-0.027	0.058	-0.058	0.054
ROCC to RIFG	0.003	0.064	-0.076	0.052	-0.131	0.044
ROCC to RIFG	0.001	0.041	-0.155	0.045	-0.075	0.041
<i>Lateral</i>						
ROCC to LOCC	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
RvOT to LvOT	0.130	0.077	-0.006	0.071	0.023	0.066
RvOT to LIFG	-0.002	0.084	-0.068	0.079	-0.100	0.071
LOCC to ROCC	0.018	0.055	-0.043	0.054	0.110	0.044
LvOT to RvOT	0.026	0.077	-0.131	0.068	0.049	0.065
LIFG to RIFG	0.000	0.083	-0.033	0.078	0.075	0.071
LIFG to RIFG	0.011	0.055	0.036	0.055	0.027	0.050
<i>Self-connections</i>						
LOCC to LOCC	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
ROCC to ROCC	0.172	0.017	0.061	0.010	-0.012	0.008
LvOT to LvOT	-0.100	0.017	-0.003	0.009	-0.050	0.008
RvOT to RvOT	-0.067	0.047	-0.089	0.024	-0.027	0.021
LIFG to LIFG	-0.020	0.045	0.004	0.023	0.076	0.020
RIFG to RIFG	-0.277	0.043	-0.034	0.020	0.033	0.022
RIFG to RIFG	0.035	0.038	0.056	0.024	0.060	0.019

(continued)

Supplementary Table 1. (continued)

<i>Feedback</i>	RW vs. PW (lexicality effect)					
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LIFG to LvOT	0.019	0.056	0.079	0.050	0.062	0.042
LvOT to LOCC	0.018	0.084	0.046	0.075	-0.125	0.062
LIFG to LOCC	-0.020	0.060	0.032	0.053	-0.033	0.045
RIFG to RvOT	0.054	0.054	-0.037	0.050	0.047	0.040
RvOT to ROCC	0.025	0.084	0.081	0.074	-0.054	0.057
RIFG to ROCC	-0.026	0.059	0.053	0.052	-0.008	0.040
<i>Feedforward</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LvOT to LIFG	0.000	0.048	-0.036	0.052	0.069	0.057
LOCC to LvOT	-0.129	0.057	-0.083	0.053	-0.118	0.045
LOCC to LIFG	-0.185	0.050	-0.033	0.039	-0.038	0.038
RvOT to RIFG	-0.017	0.049	-0.052	0.050	-0.003	0.049
ROCC to RvOT	-0.066	0.065	-0.131	0.052	-0.069	0.041
ROCC to RIFG	-0.083	0.056	0.018	0.041	-0.066	0.039
<i>Lateral</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
ROCC to LOCC	-0.071	0.082	-0.139	0.068	-0.140	0.051
RvOT to LvOT	0.025	0.084	0.032	0.077	0.049	0.063
RIFG to LIFG	-0.030	0.060	0.010	0.047	0.015	0.040
LOCC to ROCC	0.057	0.078	-0.031	0.068	-0.042	0.057
LvOT to RvOT	0.013	0.085	0.028	0.078	0.066	0.063
LIFG to RIFG	0.002	0.061	0.010	0.047	0.074	0.038
<i>Self-connections</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LOCC to LOCC	0.069	0.017	0.046	0.008	0.054	0.009
ROCC to ROCC	-0.056	0.020	0.023	0.010	0.007	0.007
LvOT to LvOT	-0.050	0.039	-0.006	0.023	0.005	0.021
RvOT to RvOT	0.046	0.038	0.041	0.022	0.030	0.020
LIFG to LIFG	0.014	0.042	0.007	0.019	0.031	0.018
RIFG to RIFG	-0.038	0.041	-0.036	0.021	-0.026	0.018

(continued)

Supplementary Table 1. (continued)

PW vs. CS (Phonological effect)						
<i>Feedback</i>	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LIFG to LvOT	-0.065	0.064	-0.028	0.049	-0.070	0.050
LvOT to LOCC	-0.071	0.087	-0.073	0.077	0.024	0.060
LIFG to LOCC	-0.035	0.057	-0.047	0.073	-0.047	0.045
RIFG to RvOT	-0.056	0.062	-0.033	0.045	0.001	0.046
RvOT to ROCC	-0.023	0.084	-0.033	0.078	0.071	0.053
RIFG to ROCC	0.037	0.056	-0.063	0.070	0.014	0.043
<i>Feedforward</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LvOT to LIFG	-0.019	0.045	-0.037	0.057	0.024	0.046
LOCC to LvOT	0.045	0.060	-0.116	0.048	0.028	0.044
LOCC to LIFG	0.188	0.053	-0.093	0.040	-0.053	0.037
RvOT to RIFG	-0.022	0.047	0.100	0.051	-0.004	0.047
ROCC to RvOT	-0.135	0.063	-0.026	0.048	-0.206	0.045
ROCC to RIFG	0.156	0.055	-0.064	0.037	-0.124	0.035
<i>Lateral</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
ROCC to LOCC	-0.142	0.080	0.121	0.066	0.173	0.055
RvOT to LvOT	-0.073	0.083	-0.008	0.076	-0.077	0.068
RIFG to LIFG	-0.043	0.055	-0.042	0.060	-0.032	0.043
LOCC to ROCC	0.051	0.081	-0.024	0.070	-0.113	0.059
LvOT to RvOT	-0.051	0.084	-0.075	0.076	-0.023	0.069
LIFG to RIFG	-0.008	0.052	-0.014	0.055	0.029	0.047
<i>Self-connections</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LOCC to LOCC	-0.132	0.022	-0.057	0.015	-0.052	0.010
ROCC to ROCC	0.012	0.021	0.052	0.013	0.028	0.011
LvOT to LvOT	0.022	0.043	0.024	0.026	-0.010	0.019
RvOT to RvOT	0.188	0.042	-0.031	0.024	0.050	0.023
LIFG to LIFG	0.042	0.046	0.064	0.024	0.044	0.016
RIFG to RIFG	-0.021	0.046	0.041	0.022	0.062	0.019

(continued)

Supplementary Table 1. (continued)

<i>Feedback</i>	CS vs. FF (Low-level letter effect)					
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LIFG to LvOT	0.004	0.063	0.103	0.049	0.057	0.048
LvOT to LOCC	0.081	0.084	0.029	0.072	0.033	0.063
LIFG to LOCC	0.013	0.047	0.046	0.048	-0.040	0.057
RIFG to RvOT	-0.018	0.065	0.004	0.051	-0.048	0.044
RvOT to ROCC	0.037	0.085	-0.120	0.075	0.042	0.058
RIFG to ROCC	-0.004	0.047	-0.032	0.047	-0.093	0.056
<i>Feedforward</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LvOT to LIFG	-0.047	0.061	0.031	0.040	-0.003	0.041
LOCC to LvOT	-0.057	0.063	0.199	0.049	0.148	0.042
LOCC to LIFG	-0.148	0.050	-0.086	0.039	0.059	0.043
RvOT to RIFG	0.009	0.066	0.017	0.039	-0.035	0.042
ROCC to RvOT	0.066	0.062	0.116	0.051	-0.050	0.044
ROCC to RIFG	-0.155	0.051	-0.012	0.037	-0.067	0.039
<i>Lateral</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
ROCC to LOCC	0.116	0.080	0.246	0.068	0.075	0.064
RvOT to LvOT	0.016	0.082	0.027	0.076	-0.030	0.070
RIFG to LIFG	0.026	0.055	-0.041	0.065	0.016	0.043
LOCC to ROCC	0.054	0.079	0.073	0.067	0.065	0.065
LvOT to RvOT	-0.054	0.086	-0.059	0.077	0.068	0.071
LIFG to RIFG	-0.014	0.057	0.017	0.067	-0.026	0.046
<i>Self-connections</i>						
	1-100 ms		1-150 ms		1-200 ms	
	Mean	SD	Mean	SD	Mean	SD
LOCC to LOCC	0.002	0.016	-0.007	0.012	-0.060	0.008
ROCC to ROCC	-0.173	0.017	-0.106	0.012	-0.149	0.011
LvOT to LvOT	0.241	0.042	-0.129	0.024	0.020	0.022
RvOT to RvOT	-0.169	0.044	0.023	0.023	0.095	0.020
LIFG to LIFG	0.093	0.036	0.054	0.020	0.056	0.022
RIFG to RIFG	-0.087	0.048	0.111	0.022	0.115	0.020

Chapter 4

Task Modulation of the Time Course of Visual Word Recognition

Yu Li, Sachiko Kinoshita, Paul Sowman, & Anne Castles

ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney

Abstract

This MEG study aimed to examine task modulation of the time course of visual word recognition. To do so, we re-analysed the data from two previous experiments using different tasks, a linguistic semantic categorisation task and a non-linguistic visual discrimination task. Both experiments used four types of critical stimulus, real words (RW), pseudowords (PW), consonant strings (CS), and false fonts (FF), to get three effects – a lexicality effect (RW vs. PW), a phonological effect (PW vs. CS), and a letter effect (CS vs. FF). Evoked field amplitudes in five time-windows after stimulus onset were extracted to examine the time course of visual word recognition in the two tasks. The electrophysiological results of root mean square and topographical analysis collectively showed that the occurrence of these three effects exhibited different patterns in the two tasks within the first 500 ms after stimulus onset; specifically, an early lexicality effect emerged in the semantic task but did not show in the discrimination task. The phonological effect emerged earlier in the discrimination task than in the semantic task. Two complementary behavioural experiments procedurally identical to the MEG experiments were carried out in two separate samples in order to obtain measures of behavioural performance for the critical conditions. Behavioural data showed that the pattern of reaction times for the four types of stimulus was $RW > PW > CS > FF$ in the semantic task. There were no significant differences among RW, PW, and CS in the discrimination task but reaction time for $CS < FF$. These findings indicate that an emphasis on high-level linguistic information sensitises the neural responses to linguistic factors, whereas an emphasis on low-level visual feature detection sensitises neural responses to surface properties. These observations collectively indicate that the time course of visual word recognition is highly dependent on task goals in which linguistic intention play a crucial role.

Introduction

Visual word recognition involves different cognitive processing stages from low-level feature detection to orthographic form analysis and then to high-level phonological and lexical-semantic access. It is a skill unique to humans, and has to be acquired through extensive experience and training. Our previous effective connectivity study has showed that early top-down feedback from the frontal cortex to ventral occipito-temporal cortex during visual word recognition is modulated by task goals (Chapter 3). Re-analysing the data from our previous experiments, we here aimed to examine how task goals modulate the time course of visual word recognition.

Behavioural investigations have established that visual word recognition is modulated by tasks with different goals (e.g., Andrews, 1982; Monsell, Doyle, & Haggard, 1989; Balota et al., 2004; Balota & Yap, 2006; Yap & Balota, 2009). For example, the lexical variable of word frequency has been found to contribute more to performance on a lexical decision task than it does on a naming task (Balota et al., et al., 2004; Yap & Balota, 2009). Behavioural findings provide direct evidence that the influences of linguistic properties on visual word recognition vary according to task goals. These behavioural investigations also reveal that task modulation of visual word recognition could take place within several hundred milliseconds before explicit responses such as a button press or naming, echoing the finding that cognitive processes during visual word recognition unfold rapidly (Sereno & Rayner, 2003; Carreiras, Armstrong, Perea, & Frost, 2014; Hauk, 2016). This brings an outstanding question - how does the task modulation of visual word recognition unfold over time? The best way to answer this question is to track the time course of visual word recognition in different tasks.

Neuroimaging techniques with high temporal resolution have advantages when it comes to tracking the time course of cognitive processes. Indeed, studies with

magnetoencephalography (MEG) and electroencephalography (EEG) have revealed influences of linguistic properties on word recognition within the first 500 ms of visual word presentation. Previous studies have shown that an effect of lexicality (differences between real words and pseudowords) takes place between 150 ms and 200 ms after stimulus onset (i.e., the N/M150 or N/M170 window; see Hauk, Davis, Ford, Pulvermuller, & Marslen-Wilson, 2006; Hauk, Coutout, Holden, & Chen, 2012), and that single word comprehension has occurred in the N/M400 window (see Kutas & Federmeier, 2011; Lau, Phillips & Poeppel, 2008; Carreiras et al., 2014; Grainger & Holcomb, 2009). These basic findings indicate that crucial cognitive processing stages of visual word recognition have been almost completed within half second after stimulus onset. Thus, electrophysiological measures with high temporal resolution can serve as an ideal tool to examine task modulation of the time course of visual word recognition (Serenio & Rayner, 2003; Pykkänen & Marantz, 2003; Pammer, 2009; Carreiras et al., 2014).

Several EEG and MEG studies have specifically examined whether task goals have an impact on the time course of word recognition (e.g., Fujimaki et al., 2009; Strijkers, Yum, Grainger, & Holcomb, 2011; Hauk et al., 2012; Chen, Davis, Pulvermuller, & Hauk, 2013, 2015; Yum, Law, Su, Lau, & Mo, 2014; Mahé, Zesiger, & Laganaro, 2015). Commonly-used tasks in these studies are lexical decision, reading aloud, silent reading, semantic categorisation and their variants. All of these tasks recruit high-level linguistic processing beyond orthographic form analysis, but have very different cognitive emphases. Lexical decision and semantic categorisation tasks specifically stress more lexical-semantic access, whereas reading aloud tasks place more emphasis on phonological retrieval and production (Balota et al., 2004; Yap & Balota, 2009). There is some evidence for the modulation of these linguistic tasks on the time course of visual word recognition. For example, Fujimaki et al. (2009) found that word-evoked neural differences between lexical and phonological decision tasks occurred mainly between 200-400 ms after stimulus onset in the left anterior temporal area. Mahé et al.

(2015) found that the event-related responses evoked by words diverged between reading aloud and lexical decision tasks from about 140 ms after stimulus onset, while early event-related responses were common to the two tasks on both waveform amplitudes and global topography. Strijkers et al. (2011) found that word-evoked event-related responses between a reading aloud task and a semantic categorisation task started to diverge at around 170 ms. These findings demonstrate that linguistic tasks with different goals modulate the time-course of visual word recognition within the first 300 ms.

All of the tasks described above are linguistic by nature, in which high-level phonological and lexical-semantic information are required to be processed before a button press or oral response occurs. Although these linguistic tasks have different goals, in addition to the processing of phonological and lexical-semantic information they all require the processing of low-level visual features and orthographic forms. Thus, comparing different linguistic tasks may not be able to detect differences in the neural responses to low-level visual features orthographic forms of visual words. Here, a non-linguistic task could be used as a potential solution because it does not require any processing of linguistic properties but only low-level visual features. A representative example of a non-linguistic task is a symbol detection task in which participants are asked to detect whether a stimulus is a hash string (#####), and no linguistic processing is necessary to perform the task (e.g., Vinckier et al., 2007). Although some fMRI studies have examined the modulation of task goals on brain activity of visual word recognition using these kinds of tasks (Yang, Wang, Shu, & Zevin, 2012; Mano et al., 2014; Yang & Zevin, 2014), they are unable to provide time course information due to the poor temporal resolution of BOLD signals. To date, only a limited number of M/EEG studies, with the required temporal resolution, have examined task modulation of the time course of visual word recognition in these two types of task - linguistic and non-linguistic (Bentin, Mouchetant-Rostaing, Giard, Echallier, & Pernier, 1999; Spironelli & Angrilli, 2007; Strijkers,

Bertrand, & Grainger, 2015; Wang & Maurer, 2017). However, the focus of these studies is somewhat limited in that each of them specifically examined task modulation of a single linguistic property and did not cover multiple time-windows, which are important for the inspection of task modulation.

Building on these previous studies, the current MEG study sought to examine task modulation of the time course of the first 500 ms visual word recognition by using a linguistic task and a non-linguistic task. To this end, we re-analysed the data of the previous two experiments using a semantic categorisation task and a visual discrimination task with the same general experiment procedure, apparatus and critical stimuli. Four types of stimulus, real words (RW), pseudowords (PW), consonant strings (CS), and false fonts (FF), were used to establish three comparisons for distinct effects, RW vs. PW (a lexicality effect), PW vs. CS (a phonological effect), and CS vs. FF (a letter effect). During the linguistic semantic categorisation task in which participants decided whether a stimulus was an animal word, high-level lexical-semantic and phonological processes must be drawn on to make a decision. However, in the non-linguistic visual discrimination task, participants were required to decide whether a visual stimulus was a hash string (e.g., ###): The participants responded by pressing a button if the stimulus was a hash string. Because no linguistic information is included in hashes, and the task could be completed with only visual features, the need for high-level linguistic processing would be expected to be minimised.

To depict the time course of lexicality, phonological, and letter effects in each task, we applied a non-parametric cluster-based permutation test (Maris & Oostenveld, 2007) to the MEG topography for five different time-windows covering the first 500 ms of evoked neural activity. This non-parametric approach is a data-driven method and solves the multiple comparison problem in EEG and MEG studies due to multiple channels and samples (Maris & Oostenveld, 2007). The three pre-defined effects as the focus of the current study helps to

reveal how the brain responds to lexical-semantic, phonological and surface properties after presentation of visual words. Based on the implications of the existing evidence that visual word recognition is highly guided by task goals, it was expected that the time course of visual word recognition would show differences across the two tasks. The nature of the tasks differs mainly in linguistic intention. The processing of both low-level visual features/orthographic form and high-level linguistic properties of visual words is required in the linguistic semantic categorisation task, whereas only low-level visual features are necessarily processed in the visual discrimination task. Thus, it was specifically hypothesised that lexicality and phonological effects on the time course of visual word recognition would be amplified in the semantic categorisation task, but would be weakened/delayed or even disappear in the visual discrimination task; in the meantime, the letter effect would be similar across two tasks.

We also conducted two complementary behavioural experiments in two different samples of adult participants. These used the same procedures as the MEG experiments and allowed us to record reaction time and accuracy for the four critical conditions (RW, PW, CS, and FF), measures which were not recorded in the two MEG experiments. Similar to the predictions above, we expected that differences between RW and PW, PW and CS, and CS and FF, especially between orthographically-legal strings RW/PW and CS/FF, would be amplified in the semantic categorisation task, but would be reduced or diminished in the visual discrimination task.

Methods²

Participants

Fifteen native English speakers (7 females, 9 males; age range: 19-28 y; mean age: 22.38 y) were recruited from the university campus to participate in the study with the semantic categorisation task. Another fifteen native English speakers (10 females, 5 males; age range: 19-31 y; mean age: 25 y) participated in the study with the visual discrimination task. These were the same participants as reported in Chapter 2 and 3, and details are provided again here for completeness, and to conform with the thesis by publication format. All participants reported that they had normal hearing, normal or corrected-to-normal vision and had no history of neurological disorders. The Edinburgh Handedness Inventory (Oldfield, 1971) was used to assess their handedness. In each experiment, fourteen participants were right-handed and one left-handed. All participants had normal reading competence measured by the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner & Rashotte, 1999). All participants had normal performance IQ as measured by the Matrices subscale of the Kaufman Brief Intelligence Test, (Second Edition; KBIT-2; Kaufman & Kaufman, 2004). See Table 1 for details. This study was approved by the Human Research Ethics Committee (Medical Sciences) at Macquarie University. A signed consent form was obtained from each participant before the experiments.

² The MEG data in Chapters 2 and 3 were re-analysed for this study.

Table 1.

Summary of behavioural tests.

	semantic categorisation task (n=15)		visual discrimination task (n=15)		<i>t</i>	<i>p</i>
	mean	SD	mean	SD		
Age (years)	22.60	2.80	23.47	3.52	-0.75	.46
Non-verbal IQ (KBIT-matrix)	118.20	8.25	113.13	8.55	1.65	.11
Word reading (TOWRE - SWE)	103.00	11.05	102.53	10.58	0.11	.91
Pseudoword reading (TOWRE - PDE)	103.47	13.08	107.40	12.49	-0.84	.41
Reading ability (combined SWE/PDE)	104.00	13.38	106.07	13.10	-0.43	.67

Overall Procedure

In each task, participants first underwent an 80-minute MEG recording and then performed the 30-minute behavioural testing which included the reading ability test, TOWRE, on the same day or the following day.

MEG Experiment Materials

The critical stimuli used in the two tasks, i.e. real word (RW), pseudoword (PW), consonant strings (CS) and false fonts (FF) were the same (see Figure 1). Within each category there were 120 exemplars divided equally into four stimulus lengths of three, four, five and six characters. All stimuli were presented in Calibri in lower case. RWs were chosen from MCWord (an Orthographic Wordform Database; for details, see <http://www.neuro.mcw.edu/mcword/>). The MCWord is a database based on the CELEX efw.cd file that contains approximately 16,600,000 examples. PWs were firstly generated based on real words through the WUGGY software (Keuleers & Bysbaert, 2010. For details, see <http://crr.ugent.be/programs-data/wuggy>) and then were carefully chosen to match RWs in

bigram frequency, trigram frequency and Coltheart's N based on MCWord database. Two-sample t-tests revealed no significant differences between RW and PW for each variable ($p < 0.10$) (see Table 1 in Chapter 2). CSs were also initially generated by the WUGGY software, with bigram/trigram frequency and Coltheart's N set at zero to keep CSs as pure letter strings without any higher-level orthographic information. As in the previous study on which this study builds (Woodhead et al., 2014), FFs were direct translations of the real words using the "Carian" font (Jane Warren, personal communication). "Carian" fonts are characters adapted from the alphabet of an obsolete Anatolian language (Melchert, 2004). To reduce possible influences of similarity to English letters, some characters whose shape was visually similar to English letters were replaced by others. Therefore, the FFs had no meanings and could not be pronounced.

Ninety-six animal words (e.g., *cat*, *dog* and *donkey*) were used as catch trials which participants were required to respond to by pressing a button in the semantic categorisation task. Ninety-six hash strings (e.g., ### and #####) were used as catch trials in the visual discrimination task (see Figure 1). 576 trials in total were included in each task.

Critical conditions				Targets	
real words	pseudowords	consonant strings	false fonts	animal words (semantic task)	hash strings (discrimination task)
guy	sog	qvr	ᵢᵢ⊕Δ	ant	###
camp	gasy	lxhp	┐⊕ᵑΠ	frog	####
blood	aloze	hhfzt	ᵑ↑ᵢᵢ□ᵑ	hippo	#####
honour	myntem	nsqrjk	□ᵢᵢᵑᵑᵑ□	cicada	#####

Figure 1. Left, examples of critical stimuli, real words, pseudowords, consonant strings, and false fonts in two experiments; *right*, examples of target stimuli in the semantic categorisation task (animal words) and visual discrimination task (hash strings).

MEG Experiment Procedure

In each task, the whole procedure included preparation, a practice run, the actual experiment and rests between blocks. The experiment was divided into six blocks. The order of stimuli was pseudo-randomised and presented on a screen in front of the participant using E-Prime 2.0 (<https://www.pstnet.com/eprime.cfm>). The distance between the screen and participants' eyes was 60 cm.

For each trial, a black screen was presented with a white cross for visual fixation; then a stimulus was presented for 500 ms, followed by a white cross which was displayed for between 2800 ms and 3400 ms. In the *semantic categorisation task*, participants were asked to attend to each trial carefully and respond with a button press when an animal name was presented on the screen. In the *visual discrimination task*, participants were asked to attend to each trial carefully and respond with a button press when an animal word was presented on the screen. Catch trials were removed from the data analysis. Before the actual experiment, a practice run was conducted to familiarise participants with the task.

MEG Data Acquisition and Pre-processing

MEG data were recorded using the KIT-Macquarie MEG160 (Model PQ1160R-N2, KIT, Kanazawa, Japan) located at Macquarie University, with participants lying in a magnetically shielded room. Data were recorded using 160 coaxial first-order gradiometers with a 50-mm baseline (Kado et al. 1999; Uehara et al. 2003). Prior to MEG data acquisition, the locations of three head landmarks (nasion and bilateral preauricular points), five marker coil positions, and the subject's head shape were recorded by 3D digitisation (Polhemus Fastrack, Colchester, VT). Participants' head position and corresponding sensor positions were

measured by energising the five marker coils briefly. The head motion was less than 6 mm for each participant in each block during the entire recording. The online sample rate was 1000 Hz.

MEG data analysis was conducted using SPM12 software (Litvak et al., 2011) implemented in Matlab 2014b (The MathWorks, Inc.). The original data sampled at 1000 Hz were high-pass filtered at 1 Hz and then low-pass filtered at 30 Hz. The filtered data were epoched from 150 ms before stimulus presentation to 600 ms after stimulus presentation. The pre-stimulus time window (-100 - 0 ms) was used for baseline correction. To reject artefacts, the Fieldtrip visual artefact rejection toolbox, which expresses every time point as a deviation from the mean over all time and channels, was used to remove extreme trials in four conditions for each participant based on the variance within each channel (http://www.fieldtriptoolbox.org/reference/ft_rejectvisual). No more than 6% of trials in any condition were excluded by the artefact rejection procedure. Subsequently, robust averaging was applied to the epoched data across trials within each condition (RW, PW, CS and FF) (Litvak et al., 2011; Wager, Keller, Lacey, & Jonides, 2005).

Event-related Field (ERF) Analysis

Prior to the topographical analysis in sensor space, root mean square (RMS) averaging across sensors was used to track the time-course of event-related fields (ERFs) and to determine global onsets/offsets of each ERF component in sensor space (Skrandies, 1990). The RMS value is the square root of the arithmetic mean of the squares of the values (averaged over trials) in all MEG sensors at a given time-point and reflects how strong an ERF is. The average RMS values for four conditions across two tasks were calculated in the first 500 ms after stimulus presentation. The ERFs peaking at around 100 ms, 150 ms, 200 ms and 400 ms are clearly evident, as can be seen in Figure 2. Thus, based on the time points of peaks of the average RMS value across the semantic categorisation task and visual discrimination task, five time-windows,

80-120 ms, 120-170 ms, 170-250 ms, 250-350 ms and 350-450 ms were obtained. The observation is generally consistent with previous studies of visual word recognition (e.g., Hauk et al., 2006; Chen et al., 2013). The RMS value in these five time-windows was averaged and used to examine differences between conditions.

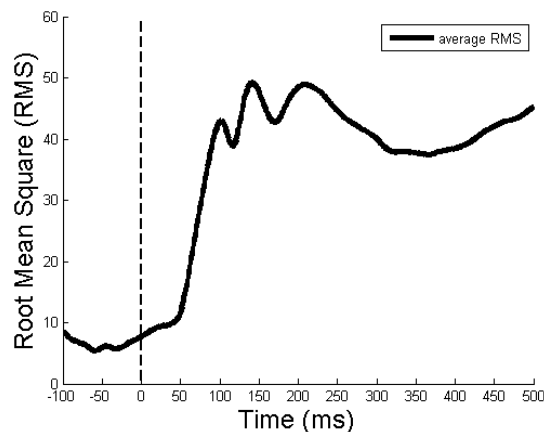


Figure 2. The average root mean square (RMS) values across the semantic categorisation task and visual discrimination task as a function of time.

Averaged MEG data was entered into a topographical analysis. To calculate the significance of differences between conditions within each task and differences between tasks, we used a non-parametric cluster-based permutation test (Maris & Oostenveld, 2007) that is implemented in the FieldTrip toolbox (Oostenveld et al., 2011). Using a clustering algorithm based on the assumption that ERF effects should be spatially clustered with adjacent sensors or channels, the permutation test estimates how many sensors show a significant effect. The approach provides a straightforward way to solve the multiple comparisons problem.

Here, we used several steps to detect the significance of differences between conditions for each of the time-windows (i.e., components). First, the topographic amplitude was averaged within time-windows pre-defined above for each sensor. Second, a dependent-sample t-test was used to calculate differences between conditions for each MEG sensor; the sensors whose significance did not exceed .05 were zeroed. Subsequently, adjacent non-zero sensors were

combined into clusters whose cluster-level t -values were the sum of all t -values within each cluster. Then, a null-distribution which assumes no difference between conditions was generated by randomly assigning participant averages to one of two conditions for 2000 times, and calculating the cluster-level statistics for each randomisation. Finally, the observed cluster-level test statistical values were compared against the null-distribution by using a Monte-Carlo estimate of the significance probabilities; if an observed value fell into one of the 2.5th percentiles of the null-distribution (i.e., two-tailed test), the observed value was considered significant. Using this approach, we tested the three comparisons pre-defined above, i.e., RW vs. PW, PW vs. CS and CS vs. FF, for each one of the five time-windows. See `ft_statfun_depsamplesT` in the FieldTrip toolbox; this function is used to test statistical significance of the differences between two dependent samples. The minimal number of neighbouring sensors was set to 5, which means that only significant clusters with at least 5 adjacent sensors would be considered as significant clusters. We used the same approach to estimate the significance of differences between the two tasks for each comparison. To do so, we firstly computed the amplitude differences between conditions within each task for each sensor and then applied the non-parametric cluster-based permutation test to the data using the function `ft_statfun_indepsamplesT` that is used to obtain statistical significance of the differences between two independent samples.

Behavioural Version of the MEG Experiments

The participants in these two tasks were asked to respond by pressing a button to an animal word or a hash string, but they did not respond to the critical conditions, RW, PW, CS, and FF. Thus, accuracy and reaction time for the critical conditions were not obtained. Because of the importance of these two behavioural measures in unveiling differences in performance between the critical conditions and how task goals modulate these differences, we conducted a

behavioural experiment for each MEG experiment in a different participant sample to collect accuracy and reaction time for each critical condition. Twenty native English speakers (18 females, 4 males; age range: 19-31 y; mean age: 21.55 y) were recruited from the university campus to participate in the behavioural experiment with the semantic categorisation task; another twenty native English speakers (19 females, 3 males; age range: 18-29 y; mean age: 20.64 y) from the university campus to participate in the behavioural experiment with the visual discrimination task. A signed consent form was obtained from each participant. Two course credits were given as compensation to each participant. Two repeated-measures ANOVAs with stimulus type as within-subject factor (four levels: RW, PW, CS and FF) were performed within each task. The significance of differences between conditions were obtained by using a Post-Hoc comparison with Bonferroni correction.

Results

Task Performance

Average RTs for the catch trials (animal words and hash strings) with correct responses in the semantic categorisation task and the visual discrimination task were 659 ms ($SD = 89$ ms) and 490 ms ($SD = 69$ ms) respectively. A two-sample t-test (two-tailed) revealed that the average reaction time for animal words in the semantic categorisation task was significantly longer than that for hash strings in the visual discrimination task ($t(28) = 5.8, p < .001$). The average accuracy for catch trials in the two tasks was 0.94 ($SD = 0.04$) and 0.98 ($SD = 0.043$) respectively. Another two-sample t-test (two-tailed) revealed that the accuracy for animal words in the semantic categorisation task was significant lower than that for hash strings in the visual discrimination task ($t(28) = -2.39, p = .02$).

Event-related Fields (ERFs)

Based on the peaks of the average RMS values across two tasks, several event-related time-windows were defined. These time-windows were M100 (80-120 ms), M150 (120-170 ms), M200 (170-250 ms), M300 (250-350 ms) and M400 (350-450 ms) (see Figure 2 for the time course of RMS values). A series of paired-t tests were conducted to detect differences between conditions to examine how tasks modulate the strength of ERFs in each window. Here the focus was on the three comparisons, RW vs. PW, PW vs. CS, and CS vs. FF, for lexicality, phonological, and letter effects respectively. The results of the semantic categorisation task showed that the difference between PW and CS was significant in the M200 and M400 ($p < .05$), and was marginally significant in the M300 ($p = .055$); the difference between CS and FF was significant in all time-windows ($p < .05$) (Figure 3). The results of the visual discrimination task showed that the difference between PW and CS was marginally significant in the M100 ($p = .089$), and the difference between CS and FF was marginally significant in the M300 ($p = .081$) (Figure 3).

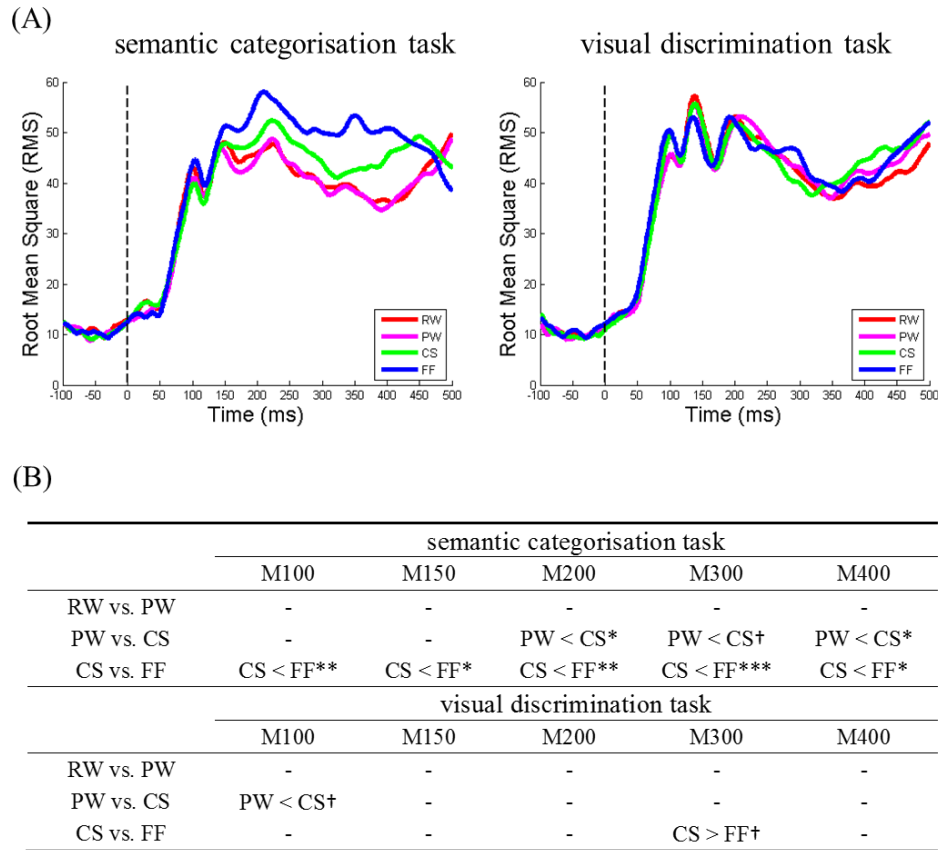


Figure 3. (A) RMS curves for four conditions in the semantic categorisation task (left) and the visual discrimination task (right); **(B)** the differences in RMS between RW and PW, PW and CS, and CS and FF in each task. RW, real words; PW, pseudowords; CS, consonant strings; FF, false fonts. † $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

See Figure 4 for the details of the topographical maps for each time-window of each condition in the semantic categorisation task (left side) and the visual discrimination task (right side). Presented below are the topographical results of three crucial comparisons, RW vs. PW, PW vs. CS, and CS vs. FF, for each task.

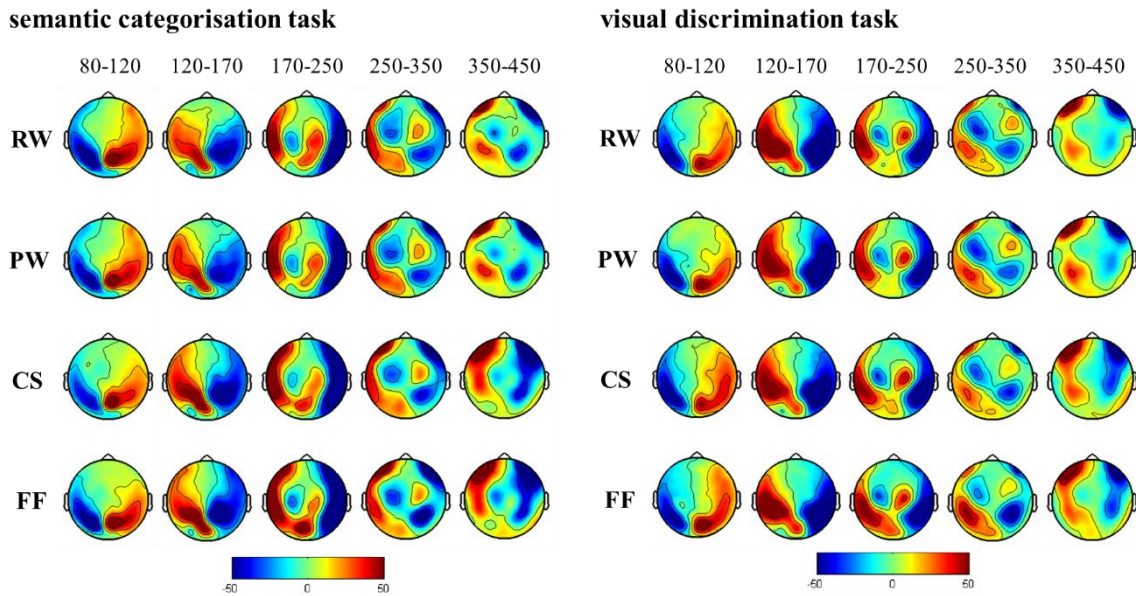


Figure 4. Topographical maps for each time-window of each condition in the semantic categorisation task (**left**) and the visual discrimination task (**right**). RW, real words; PW, pseudowords; CS, consonant strings; FF, false fonts. Colour bars denote amplitude value (fT).

Semantic categorisation task. Non-parametric cluster-based permutation tests were used to examine the lexicality effect (RW vs. PW), phonological effect (PW vs. CS) and letter effect (CS vs. FF) on the topographical activity in the M100, M150, M200, M300 and M400 time-windows. Results showed that (1) no significant clusters showing differences between RW and PW, PW and CS, and CS and FF were found in the M100 ($ps > .05$); (2) a cluster showed a significant difference between RW and PW in the M150 ($p < .05$), and the cluster was the earliest cluster showing a difference; (3) a cluster in the left hemisphere showed a significant difference between PW and CS, and a posterior cluster showed a significant difference between CS and FF in the M200 ($ps < .05$); (4) a cluster in the M300 and two clusters in the M400 showed significant differences between PW and CS ($p < .05$). See Figure 5 for details.

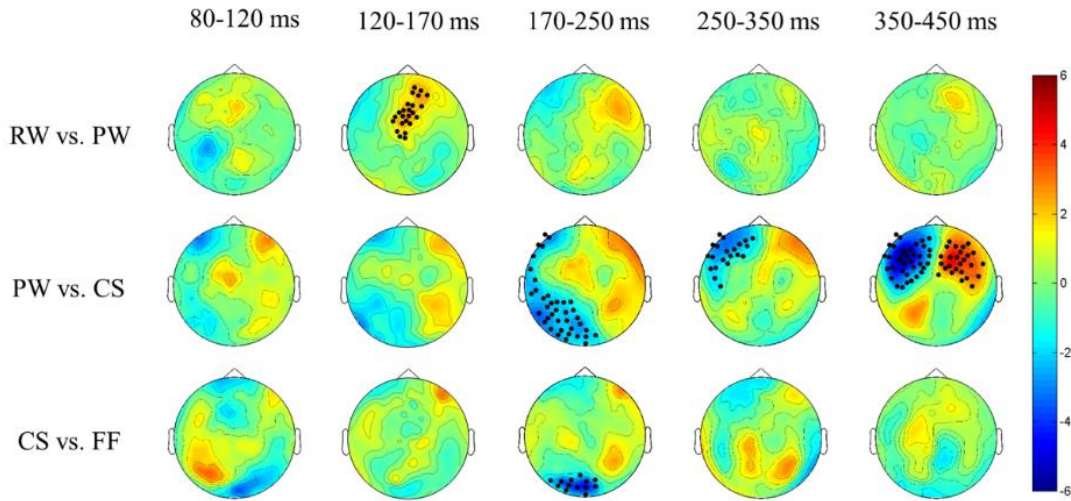
semantic categorization task

Figure 5. Topographical t-maps of non-parametric cluster-based permutation analyses for the RW vs. PW, PW vs. CS, and CS vs. FF in each time-window in the semantic categorisation task. RW, real words; PW, pseudowords; CS, consonant strings; FF, false fonts. Significant clusters ($p < .05$) were marked with black dots.

Visual discrimination task. Results showed that (1) two frontal clusters in the M100 window showed significant differences between RW and CS ($ps < .05$); (2) no clusters in the M150 and M200 showed significant differences between RW and PW, PW and CS, and CS and FF ($ps > .05$); (3) two posterior clusters in the M300 showed a significant difference between CS and FF ($ps < .05$); (4) two left clusters showed a significant difference between PW and CS, and an anterior cluster showed a significant difference between CS and FF in the M400 ($ps < .05$). See Figure 6 for details.

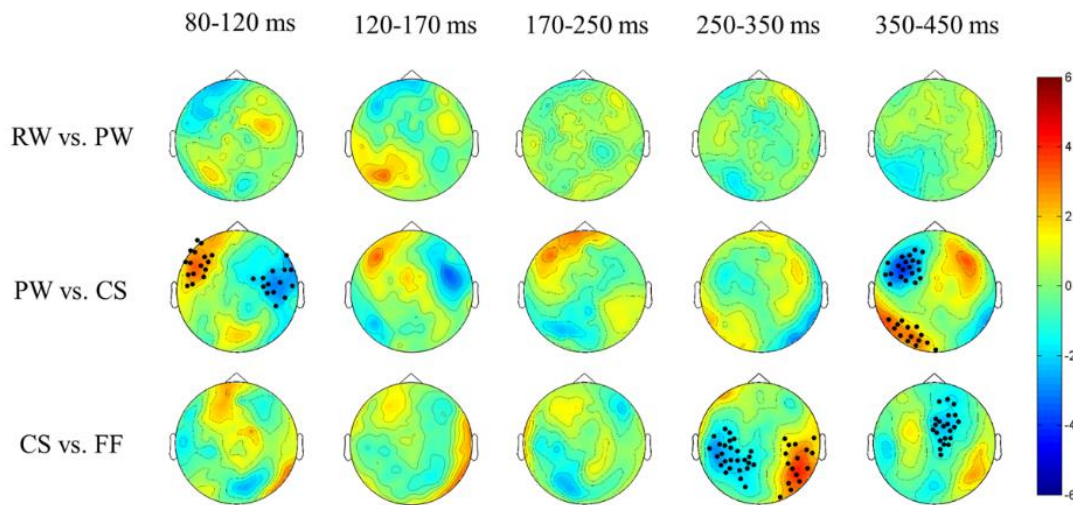
visual discrimination task

Figure 6. Topographical t-maps of non-parametric cluster-based permutation analyses for the RW vs. PW, PW vs. CS, and CS vs. FF in each time-window in the visual discrimination task. RW, real words; PW, pseudowords; CS, consonant strings; FF, false fonts. Significant clusters ($p < .05$) were marked with black dots.

Comparison between two tasks. To examine the topographical differences between tasks, a non-parametric cluster-based permutation test was used to compare differences between conditions (RW vs. PW, PW vs. CS and CS vs. FF) in the semantic categorisation task with those in the visual discrimination task. Results showed that (1) in the RW vs. PW comparison, an anterior cluster in the M150 showed a significant difference between tasks; (2) in the PW vs. CS comparison, two clusters in the M100 and a cluster in the M150 showed significant differences between tasks ($ps < .05$); (3) in the CS vs. FF comparison, a posterior cluster in the M300 and an anterior cluster in the M400 showed significant differences between tasks. See Figure 7 for details.

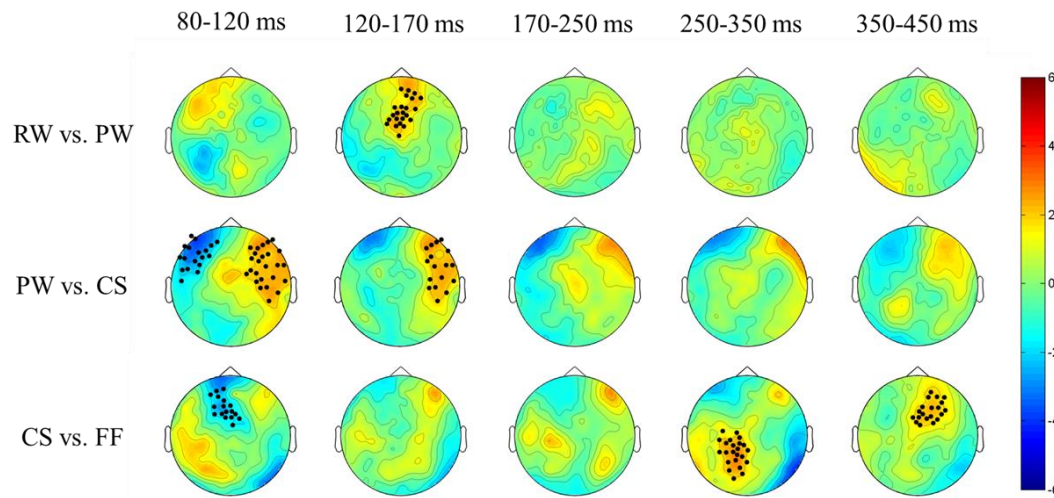
ERF differences between the two tasks

Figure 7. Topographical t-maps of non-parametric cluster-based permutation analyses between the semantic categorisation and visual discrimination tasks for the RW vs. PW, PW vs. CS, and CS vs. FF in each time-window. ERF, event-related field; RW, real words; PW, pseudowords; CS, consonant strings; FF, false fonts. Significant clusters ($p < .05$) were marked with black dots.

Behavioural Version of the MEG Experiments

Two behavioural experiments were separately conducted for each MEG experiment to obtain behavioural performance for critical conditions, RW, PW, CS, and FF. The trials with RT slower than 1400 ms and faster than 200 ms were removed. A two-sample t-test was performed to assess differences in difficulty between these two tasks. The results revealed no significant difference in the accuracy for catch trials between the semantic categorisation task ($M = 0.88$, $SD = 0.08$) and the visual discrimination task ($M = 0.91$, $SD = 0.07$) ($t(42) = -1.33$, $p = .196$), but a significant difference in the RT for catch trials between the semantic categorisation task ($M = 595$ ms, $SD = 72$ ms) and the visual discrimination task ($M = 502$ ms, $SD = 88$ ms) ($t(42) = 3.82$, $p < .001$). This indicates that it is harder to perform the semantic categorisation task than the visual discrimination task. The RTs of catch trials were similar to those obtained in the MEG experiments.

The average accuracy for all four conditions in the two tasks was high ($M > 0.97$, $SD < 0.08$), indicating that participants had a high rate of identification with very low between-subject variations (see Figure 8(A)). However, the average RTs for all four conditions in the two tasks had high between-subject variation ($453 \text{ ms} < M < 582 \text{ ms}$ and $58 \text{ ms} < SD < 101 \text{ ms}$ for the semantic categorisation task; see Figure 8(B)). Thus, the following analyses focused on the RTs. A repeated-measures ANOVA with stimulus type as the only within-subject factor was used to analyse average correct RTs for both the semantic categorisation task and the visual discrimination task. For the semantic categorisation task, the results showed that the main effect of stimulus type was significant ($F(3, 63) = 127.61$, $p < .001$). A pair-wise comparison with Bonferroni correction further revealed that the RTs for both RW and PW were significantly longer than both CS ($p < .001$) and FF ($p < .001$). One-sample t tests revealed that the difference between RW and PW was significant ($p = .021$); the difference between PW and CS was significant ($p < .001$); and the difference between CS and FF was also significant ($p = .011$). See Figure 8(B). For the visual discrimination task, the results showed that the main effect of stimulus type was significant ($F(3, 63) = 11.12$, $p < .001$). A pair-wise comparison with Bonferroni correction in multiple comparison further revealed that the RTs for both PW and CS were significantly shorter than FF ($ps < .001$). One-sample t tests revealed that the difference between RW and PW was not significant ($p = .203$); the difference between PW and CS was marginally significant ($p = .065$); and the difference between CS and FF was also significant ($p < .001$). See Figure 8(B).

A pair-wise comparison with Bonferroni correction found that the RTs for RW, PW and CS in the semantic categorisation task were longer than that in the visual discrimination task ($ps < .001$ for RW and PW, $p = .02$ for CS), and the ACCs for RW and PW in the semantic categorisation task were higher than that in the visual discrimination task ($ps < .05$). See Figure 8(B).

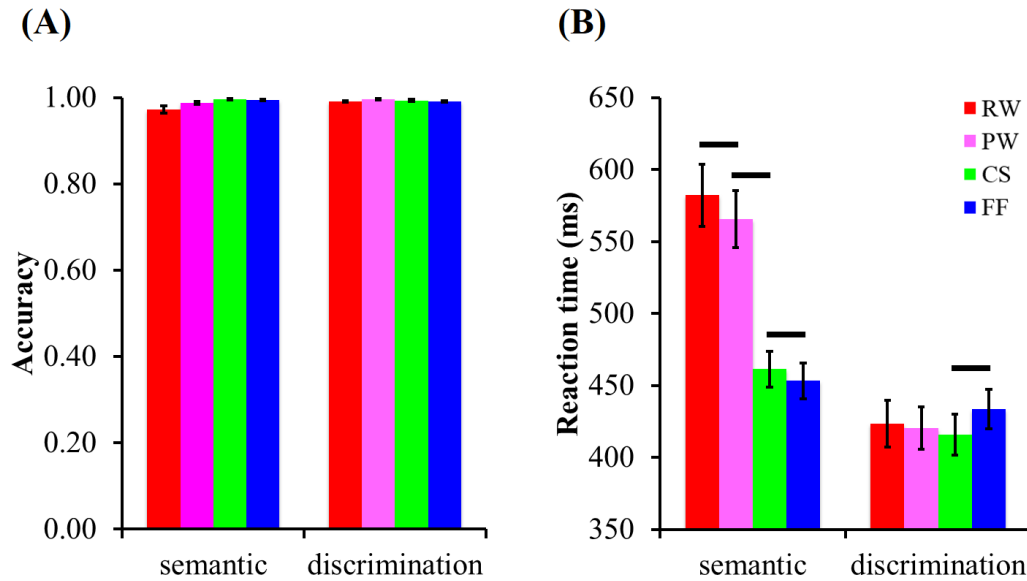


Figure 8. Average accuracy (A) and reaction time (B) for four types of stimulus in the semantic categorisation and visual discrimination tasks. Black horizontal lines denote significant differences between conditions ($p < .05$). RW, real words; PW, pseudowords; CS, consonant strings and FF, false fonts.

Discussion

Using a linguistic semantic categorisation task and a non-linguistic visual discrimination task, this study aimed to investigate how task goals modulate the time course of visual word recognition. The topographical results showed that the differences between real words and pseudowords occurred in the M150 time-window in the semantic task but disappeared in the discrimination task; the differences between pseudowords and consonant strings occurred in the M200 time-window in the semantic task, but occurred in the M100 time-window in the discrimination task; and the differences between consonant strings and false fonts occurred in the M200 time-window in the semantic task and in the M300 time-window in the discrimination task. The behavioural measures further revealed that during the semantic categorisation task, the reaction time pattern for four conditions was real words > pseudowords > consonant strings > false fonts; during the visual discrimination task, the differences in reaction

time were not significant among real words, pseudowords, and consonant strings, but the reaction times for consonant strings were shorter than for false fonts. These findings argue for a notion that the time course of visual word recognition is highly modulated by task goals, and more crucially, linguistic intention plays a key role in the time course of word-related cognitive processes.

Task Modulation of the Time Course of Visual Word Recognition

The neural differences between real word and pseudowords emerged in the M150 when the semantic task was performed, indicating an early lexicality effect. Using a lexical decision task which involves high-level lexical-semantic access, two EEG studies have previously observed early differences between real words and pseudowords (Hauk et al., 2006, 2012). However, the lexicality effect found in the semantic task was absent in all time-windows in the discrimination task. This finding is consistent with a previous observation that no significant neural difference between real words and pseudowords occurred within the first 500 ms in a non-linguistic font size same-different detection task (Bentin et al., 1999). This apparently indicates that the neural responses to differences between real words and pseudowords are modulated by task goals. Because lexical-semantic processing is not necessary and cannot help to improve the performance in the discrimination task, the neural differences between real words and pseudowords were not the cognitive focus of participants. In other words, what contributes to the absence of the lexicality effect in the discrimination task is the absence of linguistic intention. Strijkers et al. (2015) found that a key lexical factor - word frequency effect - occurring from 120 ms onward in a linguistic semantic categorisation task occurred from 220 ms onward during a non-linguistic colour categorisation task. This study indicates that the linguistic intention of a task speeds up lexical processing, while non-linguistic intention slows down the processing. The current study further showed that lexical processing is fully

diminished in a non-linguistic task. One of the crucial reasons behind this could be that the two studies focused on different aspects of lexical processing, Strijkers et al. (2015) focused on the word frequency effect while the current study focused on the overall differences between real words and pseudowords.

One particular observation did not match our expectations. Both the topographical and RMS analysis revealed that the differences between pseudowords and consonant strings occurred in the M200 in the semantic task. However, the topographical and RMS differences between pseudowords and consonant strings occurred earlier (M100) in the discrimination task. In other words, the differences between these two conditions occurred earlier in the discrimination task and was not evident in the M100 in the semantic task. These findings could be explained by the modulation of task goals. The non-linguistic discrimination task sensitised visual feature detection because the aim of the task is to detect meaningless hash strings, and facilitations from prior experience are absent. It is important to note here that pseudowords are not only pronounceable letter strings but also orthographically-legal letter strings, and consonant strings are orthographically-illegal letter strings. Therefore, it is likely that the surface distinctness between the orthographically legal and illegal strings became easier to be detected in the discrimination task. Because the discrimination task does not explicitly require the processing of orthographic legality, the early sensitivity to orthographic forms was probably an epiphenomenon of the high demands on visual feature detection. The increased sensitivity to visual features in the non-linguistic task compared to the linguistic task at early stages has been reported in a previous EEG study (Spironelli & Angrilli, 2007). They found that the early stages of brain responses reflecting visual feature extraction, the N130 (120-140 ms), evoked by words were larger in a non-linguistic visual font case judgment task compared to in linguistic phonological and semantic tasks. The increased sensitivity to low-level visual features is also supported by our previous result. That is, the early event-related field in the 50-

77 ms time-window was stronger in the non-linguistic task than the linguistic task (Figure 7 in Chapter 3). This time-window is much earlier than the M100 (80-120 ms), indicating that a task without linguistic intention but demands on low-level visual processing enhances very early stages brain activity.

A letter effect revealed by the topographical and RMS differences between consonant strings and false fonts occurred in the M200 in the semantic task. This finding is consistent with a previous study which also used a semantic categorisation task (e.g., Coch & Meade, 2016) and a study using a lexical decision task (Mahé et al., 2012). These findings indicate that the neural sensitivity to surface features appears reliably in tasks with demands on lexical-semantic access. However, the letter effect in the discrimination task appeared in a late time-window, the M300, and continued to manifest in the M400. Consonant strings, false fonts, and hash strings are all meaningless visual stimuli, and more importantly, false fonts are very similar to hash strings in terms of visual features (Figure 1). It is thereby difficult for participants to discriminate consonant strings from false fonts and then reject false fonts as non-hashes in the discrimination task. It is likely that electrophysiological signatures of the differences between consonant strings and false fonts emerged later because of high demands on low-level visual features. In an EEG study with a non-linguistic task, dot-string (e.g., ::::) detection, Appelbaum et al. (2009) found no significant differences between consonant strings and false fonts in the early 130-150 window but significant differences in the late 170-190, 240-300 and 300-500 windows. The M300 and M400 of the current study covered a long window between 250 ms and 450 ms, which is almost covered by the late 300-500 ms time-window in the aforementioned study (Appelbaum et al., 2009). The current study and Appelbaum et al. (2009) collectively indicate that the neural differences between letter strings and meaningless symbols occur late in purely visual tasks without linguistic demands. Our behavioural findings also reflected the high difficulty of rejecting false fonts as non-hashes;

that is, the reaction times for false fonts were longer than those for consonant strings (see “Task modulation of behavioural performance” below).

Briefly, these three effects in the first four time-windows (80-350 ms) in the two tasks with distinct goals suggest that task goals modulate the time course of word-related cognitive processes during visual word recognition. Because these two tasks used the same experimental procedure, critical stimuli, and apparatus, the main differences found between the two tasks could be primarily attributed to the task goal itself. Linguistic intention is included in the semantic task but not in the discrimination task. Thus, the findings tend to support that whether a goal includes linguistic intention likely affects the time-course of word-related cognitive processes such as lexicality, phonological and letter effects examined here.

M400 Effect

Topographical differences between pseudowords and consonant strings in the M400 were consistently observed in both the semantic and discrimination tasks. Past studies using a semantic task have also found this effect (e.g., Coch, 2015; Coch & Benoit, 2015; Coch & Meade, 2016; see a review by Kutas & Federmeier, 2011; Lau, Phillips, & Poeppel, 2008). This M400 effect was also found in the topographical differences between real word and consonant strings in both tasks but not found between real words and pseudowords (see supplementary Figure 1). Activity in the M400 window (N400 in EEG) is thought to reflect lexical-semantic processing and is sensitive to the orthographic legality of letter strings (e.g., Laszlo & Federmeier, 2014; Kutas & Federmeier, 2011; Deacon, Dynowska, Ritter, & Grose-Fifer, 2004). The reason why M400 is also sensitive to orthographic legality in the semantic task could be that the pseudowords may have activated semantic knowledge linked to similar looking real words (Kutas & Federmeier, 2011; Hauk, 2016) because the pseudowords were generated from real words and they still had some common aspects including similar surface

features and orthographic form neighbours (see Table 1). There is also some evidence that the M400 is sensitive to lexical neighbourhood size for both real words and pseudowords - stimuli with larger size elicit a larger effect (Holcomb, Grainger, & O'Rourke, 2002; Laszlo & Federmeier, 2011). In our case, the neighbourhood size of pseudowords and real words was well matched. Thus, it is possible that pseudowords also evoked a similar M400 effect. Thus, it is not surprising that the M400 effect was found in the semantic categorisation task due to the high demands on linguistic lexico-semantic processing.

Although the visual discrimination task relied more on low-level visual processing and required no linguistic demands, orthographic legality still triggered an M400 effect in the visual discrimination task. The reasons are still unknown. It is possible that the M400 effect for orthographic legality could not be fully removed by a non-linguistic task. A possibility behind this is that the long stimulus presentation duration (500 ms) made pseudowords be exposed to continuous eye fixation so that orthographic legality of pseudowords was sensitised, even though pseudowords could have been earlier rejected as non-words by the brain. Considering that there was also an early difference between pseudowords and consonant strings in the discrimination task, it is likely that the early and late differences between these two types of stimuli in this task probably reflects different aspects of linguistic processes. The early difference in the M100 may reflect more sensitivity to the distinctness between orthographic legality and illegality, while the late difference in the M400 may reflect partial semantic activation from pseudowords or later sensitivity to orthographic legality.

Task Modulation of Behavioural Performance

The semantic categorisation task used here entails the processing of high-level linguistic attributes of visual words, whereas the visual discrimination task does not necessarily engage this but relies more on low-level visual features. It was expected that differences in

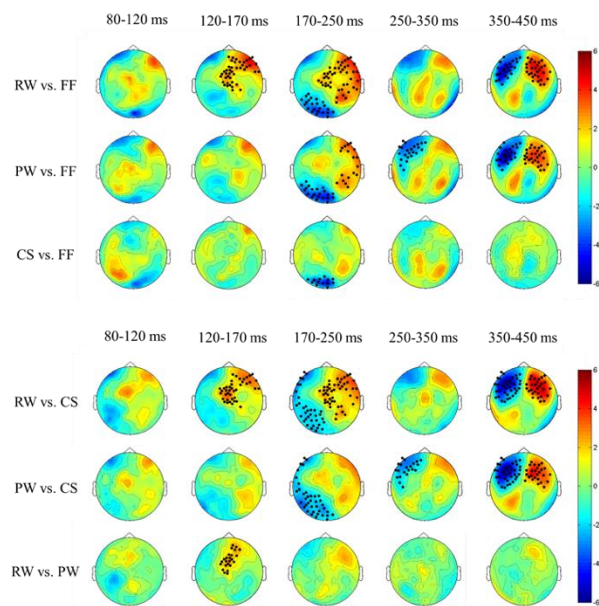
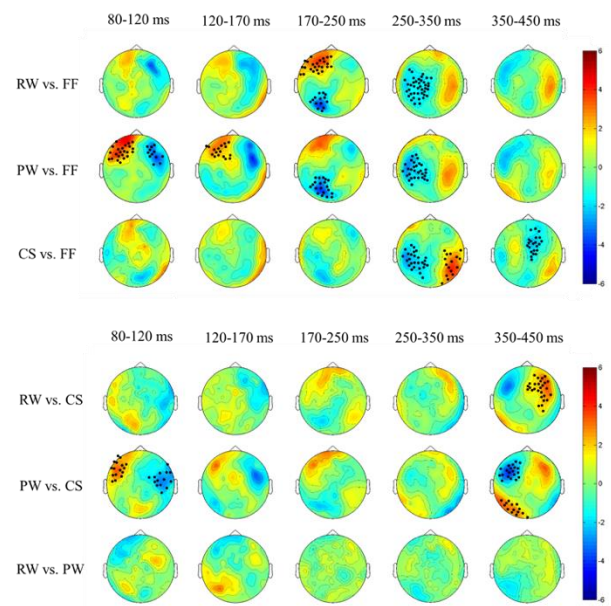
behavioural performance between different conditions would be amplified by the linguistic task, but would be reduced or diminished by the non-linguistic task. We found that the between-condition differences were indeed amplified by the semantic categorisation task, but were not all diminished by the visual discrimination task. Specifically, there were significant differences in reaction time between real words and pseudowords in the semantic task but not in the discrimination task; the reactions times for consonant strings were longer than false fonts in the semantic task while this pattern was opposite in the discrimination task (Figure 8). In addition, because real words and pseudowords both are orthographically and phonologically legal letter strings, participants took more time to reject them as non-animal words in comparison with consonant strings and false fonts during the semantic task. However, this was not the case in the discrimination task. Overall, the differences in reaction time between the two tasks suggest that task goals modulate visual word recognition in a flexible manner, which is aligned with previous suggestions (Balota & Yap, 2006; Kinoshita & Norris, 2012). The overall basic pattern of reaction time in these two tasks (Figure 8) demonstrates that linguistic intention in a task modulates the response of brain to visual words. The time course of visual word recognition that could be not revealed by behavioural index was captured by the electrophysiological measures.

Conclusion

This MEG study used a linguistic semantic categorisation task and a non-linguistic visual discrimination task to look at task modulation of the time course of visual word recognition. The results of the electrophysiological measures suggest that different types of linguistic properties - lexical-semantic, phonological/orthographic legality, and surface properties of visual words - unfold over time differently in tasks with and without linguistic demands such as lexical-semantic access. The results of the behavioural measures echo the

electrophysiological results by showing differences in behavioural performance between conditions in each task and between the two tasks. These findings collectively indicate that the first 500 ms time course of visual word recognition is guided by task goals.

Appendix A

semantic categorization task*visual discrimination task*

Supplementary Figure 1. Topographical t-maps of non-parametric cluster-based permutation analyses for all six comparisons in each time-window in the semantic categorisation task (**left**) and the visual discrimination task (**right**). RW, real words; PW, pseudowords; CS, consonant strings; FF, false fonts. Significant clusters ($p < .05$) were marked with black dots.

Chapter 5

General Discussion

Recognising a visual word involves multiple cognitive processes such as visual feature extraction, orthographic form analysis and phonological/lexical-semantic access. As reviewed in this thesis, electrophysiological investigations have shown that orthographic form analysis and phonological/lexical-semantic access occur within the first 200 ms after visual word onset, and these early neural processes are modulated by different task goals with varying linguistic demands. A robust and consistent observation in the literature is that the left inferior frontal gyrus (LIFG) is activated by visual words within the first 200 ms after stimulus onset. A recent study further found that this region sends early top-down feedback to the left ventral occipito-temporal cortex (LvOT) during visual word recognition (Woodhead et al., 2014). Building on this previous work, this thesis used magnetoencephalography (MEG) to investigate the nature of this early top-down feedback from LIFG to LvOT (Chapter 2) and to explore how task goals modulate it (Chapter 3). This thesis further used MEG data from these experiments to examine task modulation of the time course of visual word recognition (Chapter 4).

In this chapter, I will summarise the main findings in each empirical chapter and discuss how these findings contribute to the understanding of the early neural dynamics of visual word recognition. I will first focus on the contribution of this thesis to the existing neural models, especially the Interactive Account of the LvOT. Then I will move to a discussion of the task-dependence and -independence of visual word recognition. I will discuss the importance of examining early functional integration of visual word recognition and the appropriateness of different tasks in examining the neural correlates of this process. Finally, I will discuss the limitations of this thesis and potential future directions before making a general conclusion.

Summary of Main Findings

Extending a previous study revealing that the LIFG-to-LvOT top-down feedback is stronger for words compared with visual symbols (false fonts) at the early stages of visual word recognition (Woodhead et al., 2014), **Chapter 2** aimed to determine the nature of this early top-down feedback. To this end, I conducted an MEG study using a semantic categorisation task and applied dynamic causal modeling (DCM), measuring directional influences of one region on another to estimate how this top-down feedback reflects lexicity effects (real words vs. pseudowords), phonological effects (pseudowords vs. consonant strings) and letter effects (consonant strings vs. false fonts). It was expected that lexicity and phonological effects but not letter effects would influence the early top-down feedback, because the LIFG is thought to involve high-level phonological and lexical-semantic processing. It was shown that the LIFG-to-LvOT top-down feedback was stronger for real words than for false fonts in the 1-150 ms time-window, replicating Woodhead et al. (2014). More importantly, the top-down feedback was stronger for real words than for pseudowords in both 1-150 ms and 1-200 ms time-windows, indicating a lexicity effect; the feedback was stronger for pseudowords than for consonant strings in the 1-200 ms time-window, indicating a phonological effect; furthermore, the feedback was also stronger for consonant strings than for false fonts both 1-150 ms and 1-200 ms time-windows, indicating a letter effect. These findings collectively suggest that compared with phonological information, lexical-semantic and letter information have an earlier influence on top-down feedback. All findings of this study point to early top-down feedback not being purely triggered by a single property of visual words, but by multiple properties; and that lexical-semantic and letter properties trigger top-down feedback earlier than phonological properties.

Following the study reported in Chapter 2, **Chapter 3** aimed to investigate whether the LIFG-to-LvOT top-down feedback is modulated by task goals. I conducted an MEG experiment using a visual discrimination task requiring participants to detect symbol strings (a string of # signs) and applied DCM for evoked responses to the data. On the assumption that the visual discrimination task relies on low-level visual feature processing but not necessarily high-level linguistic properties, it was expected that the early top-down feedback in the discrimination task would be different from that in the semantic task in Chapter 2. The results showed that the LIFG-to-LvOT feedback was comparable for real words compared to false fonts in all time-windows in the visual discrimination task, which is in contrast to the findings in the semantic task in Chapter 2. Interestingly, the LIFG-to-LvOT feedback was stronger for real words than for pseudowords in both 1-150 ms and 1-200 ms time-windows in the visual discrimination task, which was similar to that in the semantic task, indicating a task-independent lexicity effect. However, in the discrimination task the LIFG-to-LvOT feedback was stronger for consonant strings than for pseudowords in the 1-200 ms time-window and was stronger for false fonts than for consonant strings in both 1-150 ms and 1-200 ms time-windows, which was almost opposite to that in the semantic task, indicating task-dependent phonological and letter effects. Similar to the semantic task, compared with phonological information, lexical-semantic and letter information have an earlier influence on this top-down feedback. These results reveal that several aspects of the early top-down feedback from frontal cortex to ventral occipito-temporal cortex are modulated by task goals, but that influence of lexical-semantic properties appears to occur independently of task goals.

Using the MEG data from Chapters 2 and 3 but adopting a different focus, **Chapter 4** sought to examine task modulation of the time course of visual word recognition in the linguistic semantic categorisation task and the non-linguistic visual discrimination task. Lexicality effects (defined by real words vs. pseudowords), phonological effects (defined by

pseudowords vs. consonant strings), and letter effects (defined by consonant strings vs. false fonts) were examined in five time-windows covering the first 500 ms after stimulus onset. Because the semantic categorisation task recruits lexical-semantic properties of visual words and the visual discrimination task only requires processing of visual features but not necessarily any linguistic information, it was hypothesised that the time course of visual word recognition would be modulated by the two different tasks. The results of event-related field analysis showed that a lexicality effect emerged in the M150 for the semantic task, but was absent in all windows in the discrimination task; a phonological effect emerged in the M200 in the semantic task and in the M100 in the discrimination task; a letter effect emerged in the M200 in the semantic task and in the M300 in the discrimination task. The complementary behavioural studies showed that the pattern of reaction times of four types of stimuli was real word > pseudowords > consonant strings > false fonts in the semantic task. For the visual discrimination task, there were no significant differences among real word, pseudowords, and consonant strings, but the reactions times for consonant strings were shorter than for false fonts. In addition, overall reaction times were much shorter in the discrimination task than the semantic task. This study demonstrated that the time course of visual word recognition unfolds differently in a non-linguistic task and a linguistic task and in general is modulated by task goals.

The findings revealed by this thesis have several implications for the understanding of the early neural dynamics of visual word recognition. First and foremost, they add new knowledge to inform neural models of visual word recognition.

Neural Models of Visual Word Recognition

The findings in Chapters 2 and 3 contribute to the current understanding of neural models of visual word recognition. According to the neural model proposed by Price and Mechelli (2005), the reading brain has different systems which have subregions with different functional roles. Within the LIFG, the anterior and ventral part is more responsible for semantic processing while the posterior and dorsal part is more responsible for phonological processing. The current thesis did not divide this frontal language region into subregions, but did find that lexical-semantic and phonological properties of words could influence the early LIFG-to-LvOT top-down feedback. This observation is consistent with the proposal that the LIFG is a region for lexical-semantic and phonological processing. In addition, it was found that letter effects also have an influence on this early top-down feedback, indicating the LIFG is also involved in the processing of letter information about visual words. This is supported by previous fMRI investigations in which the LIFG was found to be activated by single letters (Flowers et al., 2004; Liu et al., 2010, 2011). Thus, the current findings tend to support that the LIFG is a multi-functional region for visual word recognition.

There are two models with differing views regarding the functional roles of the LvOT in reading and visual word recognition. According to the local combination detector (LCD) model proposed by Dehaene, Cohen, Sigman, & Vinciker (2005), the visual word form area (VWFA) in LvOT is specific for written words. It computes visually abstract, pre-lexical orthographic representations in a primarily feed-forward manner. In contrast, according to the Interactive Account proposed by Price & Devlin (2011), the LvOT is not specific for written words, but is also responsible for encoding other visual stimuli such as pictures. It functions as an interface linking bottom-up, feedforward, low-level visual information from occipital cortex, and top-down feedback of high-level linguistic information from language areas. Chapters 2

and 3 demonstrate that the LvOT receives top-down feedback from high-order language areas at the early stages of visual word recognition, and this early top-down feedback is influenced by high-level lexical-semantic and phonological properties and low-level letter information. Thus, the results described in these two chapters can be well explained in the framework of the Interactive Account (Price & Devlin, 2011). However, the findings do not exclusively speak against the view that the VWFA within the LvOT functions as an area for orthographic form analysis (Dehaene et al., 2005; Dehaene & Cohen, 2011).

Both the LCD model and the Interactive Account hold the view that the lateralisation of the spoken language network influences lateralisation of the vOT. In an EEG study, Cai, Lavidor, Brysbaert, Paulignan, & Nazir (2008) previously observed that frontal lobe activity during a word generation task co-lateralised with the occipito-temporal activity during a word reading task, which demonstrates the top-down influences of language network in the development of the vOT. A crucial hub in the spoken language network is the LIFG. The current thesis revealed that early top-down feedback from frontal to ventral occipito-temporal cortex showed a strong left-lateralisation in both the semantic categorisation task and the visual discrimination task. It demonstrates that early top-down feedback is highly influenced by the left-lateralised spoken language network. The left-lateralisation of early top-down feedback is generally in line with both the LCD model and the Interactive Account in that the lateralisation of language network determines the lateralisation of the vOT.

The occurrence of the left-lateralised early top-down feedback for lexicality and phonological effects in both linguistic and non-linguistic tasks further indicates that the left-lateralisation of early top-down feedback is highly robust and is free from the modulation of task goals. Left-lateralisation of early top-down feedback has also been revealed by Woodhead et al. (2014) who used a semantic categorisation task. Previous studies with similar linguistic tasks also found left-lateralised word-recognition network (see a meta-analysis by McNorgan,

Chabal, O'Young, Lukic, & Booth, 2015). Previous fMRI studies with non-linguistic tasks such as hash string detection and size judgment have also observed similar left-lateralised activity of the vOT and IFG (e.g., Vinckier et al., 2007; Liu et al., 2008). For example, Liu et al. (2008) found that the activity in the IFG and vOT elicited by Chinese characters was left-lateralised in a font size detection task. Using a hash detection task, Vinckier et al. (2007) found a neural hierarchy in the left vOT with increasing sensitivity to word-like stimuli progressing from the posterior part to the anterior part, but not in the right vOT. A novel finding of Chapters 2 and 3 is the demonstrated left-lateralised early top-down feedback that appears in both a linguistic semantic categorisation task and a non-linguistic visual discrimination tasks.

In summary, this thesis extends the current Interactive Account (Price & Devlin, 2011) by adding new knowledge that the LvOT receives top-down feedback modulation from the frontal language area, even at the very early stages of visual word recognition. More importantly, lexical-semantic, phonological, and letter information all influence early top-down feedback from frontal to ventral occipito-temporal cortex at early stages; the left-lateralised top-down feedback occurs in the first 200 ms of visual word recognition.

Task-Dependence of Early Neural Activity

Behavioural studies have revealed that task goals modulate visual word recognition (Balota et al., 2004; Balota & Yap, 2006; Yap & Balota, 2009). There is also a growing body of neuroimaging evidence that brain activity during visual word recognition, including local activation and inter-regional connectivity, is dependent upon task goals (e.g., McNorgan et al., 2015; Bitan et al., 2005, 2006). This behavioural and neuroimaging evidence demonstrates the high flexibility of the brain in responding to visual words.

In this thesis, task dependence of the early neural dynamics of visual word recognition was examined. Chapters 3 and 4 used the same dataset but different methods. Task-dependent brain activity was found in both Chapters 3 and 4. Interestingly, the overall strength of the event-related fields at very early stages, in the 50-80 ms window, was modulated by task goals. The combined findings of task modulation from these two chapters demonstrate that a task with specific goals biases brain activity at the early stages of visual word recognition. However, not all aspects of neural dynamics are modulated by task. For example, the findings in Chapter 3 indicate that the influences of the lexical-semantic properties of visual words on early top-down feedback are independent of task goals. The finding in Chapter 4 that the M400 differences in the neural responses between pseudowords and consonant strings occurred in both tasks (Figures 6 and 7), demonstrates that M400 for orthographic legality may be free from the modulation of task goals. These findings imply that the human brain flexibly responds to visual words in serving specific cognitive goals in different tasks, but some processes may not be fully guided by task goals. Due to the earliness of task modulation effects observed in this thesis, these findings extend existing observations from behavioural and fMRI studies examining task modulation of cognitive and neural aspects of visual words. With the focus on task modulation of early stage inter-regional connectivity, it builds on previous electrophysiological studies examining task modulation of regional activity during visual word recognition (e.g., Chen, Davis, Pulvermuller, & Hauk, 2013, 2015; Strijkers, Bertrand, & Grainger, 2015; Wang & Maurer, 2017).

It is important to note that Chapter 4 used the data from previous two chapters, but focused on different neural measures: Chapter 4 on time course while Chapters 2 and 3 on inter-regional connectivity. However, there were some apparent discrepancies in different effects across these different investigations. For example, a lexicality effect for early top-down feedback was found in both the semantic and discrimination tasks, but it was not found in the

time course measure of the discrimination task. The reason for these discrepancies remains to be explored. Chapters 2 and 3 examined inter-regional connectivity, which reflects neural interaction or communication between two local regions with different functions *per se*. However, the focus of Chapter 4 was on the event-related field activity that measures summed electromagnetic activities on the scalp. Using different neural measures could have led to the observed discrepancies. Using the discrepancies as a starting point, future studies could examine the relationship between these two different neural measures.

Early Stage Functional Integration

Chapters 2 and 3 further highlight the importance of early stage functional integration during visual word recognition. The understanding of the neural correlates of visual word recognition over the past 150 years can be generally divided into two parts. The first part focuses on *functional segregation*, i.e., local activation and corresponding functions. For example, investigations with brain damaged patients originally found that the LIFG and LvOT play an important role in speech production and visual word recognition respectively (Broca, 1861; Dejerine, 1892). Using modern neuroimaging techniques, especially fMRI, researchers have further found that visual word recognition involves multiple local areas in the brain, including the occipital cortex, LvOT, LIFG, and some subcortical structures such as the putamen, caudate and thalamus (e.g., Pugh et al., 2000; Bolger, Perfetti, & Schneider, 2005; Houdé, Rossi, Lubin, & Joliot, 2010; Taylor, Rastle, & Davis, 2013; Martin, Schurz, Kronbichler, & Richlan, 2015). As neuroimaging methods have developed, researchers have increasingly examined how remote brain areas interact during reading-related tasks. This is *functional integration*, and involves considering neural interactions between remote regions with different functional roles, which can be quantified by neural measures such as functional and effective connectivity (Friston, 2011). The application of functional and effective

connectivity to neuroimaging data has helped reveal neural couplings between reading-related regions such as the LIFG and LvOT (van der Mark et al., 2011; Finn et al., 2013; Bitan et al., 2005, 2006; Heim et al., 2009; Perrone-Bertolotti, Kaufmann, Pichat, Vidal, & Baciú, 2017) and their predictive roles in reading competence (Howitz, Rumsey, & Donohue, 1998; van der Mark et al., 2011; Koyama et al., 2011; Finn et al., 2013; Li et al., 2017). These inter-regional connections are structurally supported by white matter pathways (Ben-Shachar, Dougherty, & Wandell, 2007; Yeatman, Rauschecker, & Wandell, 2013; Vandermosten, Boets, Wouters, & Ghesquiere, 2012).

Studies of functional segregation and integration are complementary to each other in contributing to a better understanding of how visual words are processed in the brain. Because cognitive processes in visual word recognition take place so rapidly (Serenó & Rayner, 2003; Carreiras, Armstrong, Perea, & Frost, 2014; Hauk, 2016) and a word can be recognised within several hundred milliseconds (Keuleers, Diependaele, & Brysbaert, 2010; Keuleers, Lacey, Rastle, & Brysbaert, 2012), studies exploring this must use neuroimaging measures with high temporal resolution. As Hauk states, “If the brain processes we are interested in are fast, then our measurements should be fast.” (Hauk, 2016, p. 1073). In the current thesis, with the application of DCM for evoked responses to MEG data, Chapters 2 and 3 examined directional interactions from LIFG to LvOT in the first 200 ms of visual word recognition. The results of both Chapters 2 and 3 indicate that although the influences of different properties of words on the early top-down feedback vary, the neural interactions between the LIFG and the LvOT occur at the early stages of visual word recognition. This type of investigation is beyond the scope of fMRI techniques due to their relatively poor temporal resolution.

These new observations in Chapters 2 and 3 update our previous understanding of neural interactions during visual word recognition in that they add new knowledge about early inter-regional effective connectivity. The field calls for more research to uncover early stage

functional integration in the brain during reading and visual word recognition. This thesis may provide a basis for future research on this topic.

Appropriateness of Word-Recognition Tasks

In the field of reading and visual word recognition, researchers usually use word reading, lexical decision, semantic categorisation tasks, and their derived versions to investigate how the brain responds to visual words (e.g., McNorgan et al., 2015; Taylor et al., 2013; Woodhead et al., 2014). These linguistic tasks have different emphases on linguistic properties, either phonologically oriented or semantically oriented. A smaller number of studies have used non-linguistic tasks such as feature detection, repetition detection, symbol detection, colour categorisation tasks (e.g., Price, Wise, & Frackowiak, 1996; Binder et al., 2006; Vinkier et al., 2007; Liu et al., 2008). Some researchers term linguistic tasks as explicit tasks because they all by default require cognitive processing of linguistic properties of visual words, and term non-linguistic tasks as implicit tasks because they all by default do not explicitly require word-recognition processing (Price et al., 1996; Liu et al., 2008). Implicit tasks, compared with explicit tasks, are thought to reduce phonological and lexical-semantic processing and restrict as much as possible attentional and top-down effects on brain activity (Price et al., 1996; Binder, Medler, Westbury, Liebenthal, & Buchanan, 2006; Vinkier et al., 2007; Liu et al., 2008). For this reason, they are appropriate to examine sensory-driven brain activity evoked by visual words. Implicit tasks have been applied to investigations of the LvOT during visual word recognition and reading (Price et al., 1996; Binder et al., 2006; Vinkier et al., 2007; Liu et al., 2008; Vartiainen, Liljeström, Koskinen, Renvall, & Salmelin, 2011; Strijkers et al., 2015), even though not all studies label these tasks as implicit tasks.

In the current case, the semantic categorisation task required high-level lexical-semantic access while the visual discrimination task only required low-level visual feature detection. If the classification above is used here, the semantic task is explicit and the discrimination task is implicit. Note that the discrimination task is similar to the symbol detection task used in Vinkier et al. (2007) in that these authors also used occasionally-presented hash strings as targets. The findings of effective connectivity, time course, and reaction time in chapters 3 and 4 collectively demonstrate that the visual discrimination task biases participants' attention to the detection of low-level features and also evokes early top-down feedback from frontal to ventral occipito-temporal cortex. A good example is that the reaction times for false fonts were shorter than consonant strings during the discrimination task, but this pattern reversed in the semantic task; and the differences in reaction time between real words and pseudowords were almost absent in the discrimination task, but were present in the semantic task. Interestingly, the discrimination task appeared to enhance early neural sensitivity to orthographic legality and overall enhanced the strength of early event-related fields. This general pattern is supported by a previous EEG study using a non-linguistic task (implicit) and two linguistic tasks (explicit) (Spironelli & Angrilli, 2007), which also found that, compared with the linguistic tasks, the non-linguistic task led to increased brain activity for words in visual cortex.

An open question is whether brain activity during an “implicit” task is purely sensory-driven and bottom-up by nature. Specifically, for the LvOT, does the activity within this region during an “implicit” task fully reflect sensory-driven processing? The findings in the discrimination task in Chapter 3 indicate that although the discrimination biased participants' attention to low-level visual features, the LvOT also receives early top-down feedback from the LIFG; thus, there is no fully sensory-driven activity during an “implicit” task. The task-modulation of brain activity during visual word recognition in this thesis does not speak against

the classification of explicit versus implicit. However, this thesis tends to support that the nature of the task used should be carefully considered in interpreting findings. In addition, it is important to consider the cognitive biases produced by a task when interpreting brain activity during that task. McNorgan et al. (2015) provides a good example. That study examined task by lexicality interactions in the brain, and explicitly stated that “interpreting lexicality effects should account for task” (McNorgan et al., 2015, p. 157).

The discussion above also raises a further question: what a specific task should one use to examine neural activity during visual word recognition? The viewpoint here is that the answer here depends on what specific cognitive processes of visual word recognition one wants to focus on. If one aims to examine neural responses of phonological processing, then a phonological-oriented task such as word reading and rhyming tasks which emphasise phonological retrieval and manipulation can be used. In contrast, if the focus is on neural responses associated with accessing lexical representations and the meanings of words, then a more semantically-based task might be chosen.

Limitations and Future Directions

Although the studies presented in this thesis have contributed new knowledge about the early neural dynamics of visual word recognition, they have some limitations. As discussed in Chapters 2 and 3, not all DCM models in the six-node network were established due to computational limitations, which could lead to a possibility that the estimation of information flow in the network was biased. Future studies could use powerful computation to estimate all possible models in a pre-defined word-recognition network, as Ge et al. (2015) did in a recent fMRI study of language comprehension. In addition, the two DCM chapters specifically focused on the first 200 ms of top-down feedback, but not late time-windows such as 200-500

ms after stimulus onset. Obviously, only the first 200 ms brain activity cannot fully capture the full picture of visual word recognition. There is mounting evidence that late time-windows involve some key visual word recognition cognitive processes such as word comprehension, stimulus re-evaluation, and episodic memory (Carreiras et al., 2014; Hauk, 2016). Building on the effective connectivity findings in this thesis, future studies could examine how information exchange in the brain, especially top-down feedback, occurs at late stages of visual word recognition.

The findings of this thesis add to a better understanding of the early stages word recognition. Several future directions can be suggested from the current work. The first pertains to the possible functional relevance of neural oscillations in early top-down feedback. Information exchange in the brain is not limited to neural interactions in terms of time-series coupling between local regions. Previous studies have also established that neural oscillations play an important role in transferring information between remote local regions (Fries, 2015). Intriguingly, recent electrophysiological investigations have revealed an overall pattern that gamma neural oscillations carry bottom-up feedforward information flow from lower-order to higher-order regions and beta neural oscillations carry top-down feedback information flow from higher-order to lower-order regions in macaque and human visual cortex (Bastos et al., 2015; Michalareas et al., 2016). In the field of reading, Schoffelen et al. (2017) found that during sentence reading, directed connections from temporal to inferior frontal regions peaked in the alpha-band frequency while directed connections from inferior frontal and to temporal regions peaked in the beta-band. These findings align with previous observations of differences in frequency between feedforward and feedback connections in the visual cortex (Bastos et al., 2015; Michalareas et al., 2016). Kujala et al. (2007) found that the LvOT is an important hub sending information to other regions at a frequency of 8-13 Hz during word-by-word text

reading. These studies suggest that frequency-bands of neural oscillations are likely involved in information exchange in the reading brain.

Applying DCM for evoked responses to MEG data, Chapters 2 and 3 has revealed that the LIFG sends early top-down feedback to the LvOT during visual word recognition and this feedback is modulated by task goals. An open question to be answered is that what specific frequencies play an important role in sending early top-down feedback and bottom-up feedforward in the reading network. Previous MEG studies have revealed that the early activation of LIFG during visual word recognition is mainly driven by neural oscillations in beta band such as 10-20 Hz (Pammer et al., 2004), 10-25 Hz (Cornelissen et al., 2009), 15-35 Hz (Wheat, Cornelissen, Frost, & Hansen, 2010), and 15-20 Hz (Klein et al., 2014). These studies indicate a possibility that beta band oscillations in the LIFG may function as top-down feedback. But this possibility needs to be confirmed. As such, at what frequency bands neural oscillations function bottom-up feedforward in the reading network is still unexamined. A further question is how these two types of neural oscillations interact with each other to produce efficient and effective word recognition. Using effective connectivity techniques, such as Granger causality analysis or DCM for induced responses, future work could contribute to this issue.

The second direction pertains to early top-down feedback in different linguistic tasks. Two very different tasks, a linguistic semantic categorisation task and a non-linguistic discrimination task, were used in this thesis. Chapter 3 specifically revealed a strong task modulation on early top-down feedback by showing differences between the two tasks in three effects (lexicality, phonological, and letter effects) on the LIFG-to-LvOT feedback. This reveals that task goals modulate early top-down feedback; however, how early top-down feedback occurs in different linguistic tasks is still unclear. A linguistic task can be oriented towards phonology, lexicality, or semantics. A good example task for each of these tasks is

word reading (reading aloud or silent reading), lexical decision, and semantic decision/categorisation. The time course of word recognition has been previously examined in studies with different linguistic tasks (e.g., Chen et al., 2013, 2015; Mahé et al., 2015). The current results have showed that task modulation of early top-down feedback takes place even within 200 ms after stimulus presentation. Because the LIFG has multiple functions in visual word recognition (Price & Mechelli, 2005; Price, 2012), it would be intriguing to examine the degree to which different kinds of linguistic task modulate this early top-down feedback. For example, does this feedback takes place differently in a lexically-stressed lexical decision task and a phonologically-stressed word reading task? Future studies with more than one linguistic task could address these questions.

The third future direction concerns the functional roles of other word-recognition nodes in information exchange in the brain. Chapters 2 and 3 of this thesis focused on early top-down feedback from LIFG to LvOT. However, reading-related functions are not limited to these two areas but also include the left posterior superior temporal cortex, anterior temporal lobe (Pugh et al., 2000; Price & Michelle, 2005; Price, 2012; Taylor et al., 2013). The left posterior superior temporal cortex plays an important role in auditory word form processing including grapheme-to-phoneme conversion (Pugh et al., 2000; Taylor et al., 2013). The anterior temporal lobe is thought to be a hub for semantic processing (Ralph, Jefferies, Patterson, & Rogers, 2016). These regions are regarded as crucial nodes for visual word recognition and reading and they are involved in the processing of different linguistic properties and interact with each other during visual word recognition (Hoffman, Ralph, & Woollams, 2015; Boets et al., 2013; Finn et al., 2013; Hancock, Richlan & Hoefft, 2017). However, they were not included in the word-recognition network in Chapters 2 and 3 because the focus of these two chapters was specifically on top-down feedback from frontal to ventral occipito-temporal cortex. Thus, it is still unclear what potential roles these regions play and how they interact with the LIFG

and LvOT at the early stages of visual word recognition. It would be interesting for future studies to examine these outstanding questions. For example, the question of how the left anterior temporal lobe is involved in lexical-semantic access at the early stages of visual word recognition, and how this region communicates with the LIFG/LvOT, remains to be investigated. Hoffman et al. (2015) found neural interactions between these regions, but could not reveal early neural interactions due to the poor temporal resolution of BOLD signals. What is more, some areas have subregions. For example, LIFG has three subregions that are thought to have different preferential functions; the more dorsal and posterior part is involved in phonological processing while the more ventral and anterior part is involved in lexico-semantic processing (Price & Michelle, 2005). Future studies could pay more attention to these subregions in feedback and feedforward information exchange for visual word recognition. These three directions for future studies are important for the exploration of early neural interactions in the brain when a visual word is recognised.

Conclusion

The current thesis examined the early neural dynamics of visual word recognition. The results indicate that early top-down feedback from the LIFG to LvOT is influenced by multiple linguistic properties of visual words, such as lexical-semantic and phonological properties. The results also indicate that the influences of linguistic properties on this early top-down feedback are modulated by task goals, but lexical-semantic influences are free from task modulation. Furthermore, the time course of visual word recognition within the first 500 ms after stimulus onset is guided by task goals, in which linguistic intention is an important contributor. The combination of two neural measures, and the examination of effective connectivity between areas and time course, add to the extant knowledge of the early neural dynamics of visual word recognition. Future research can use the findings of this thesis as a starting point to investigate related topics, such as the role of neural oscillations and the possible functions of other key word-recognition regions in early top-down feedback.

References

- Alcauter, S., Garcia-Mondragon, L., Gracia-Tabuenca, Z., Moreno, M. B., Ortiz, J. J., & Barrios, F. A. (2017). Resting state functional connectivity of the anterior striatum and prefrontal cortex predicts reading performance in school-age children. *Brain and Language*, 174, 94-102. doi:10.1016/j.bandl.2017.07.007
- Andrews, S. (1982). Phonological Recoding - Is the Regularity Effect Consistent. *Memory and Cognition*, 10, 565-575. doi:10.3758/Bf03202439
- Appelbaum, L. G., Liotti, M., Perez, R., Fox, S. P., & Woldorff, M. G. (2009). The temporal dynamics of implicit processing of non-letter, letter, and word-forms in the human visual cortex. *Frontiers in Human Neuroscience*, 3, 56. doi:10.3389/neuro.09.056.2009
- Araujo, S., Faisca, L., Bramao, I., Reis, A., & Petersson, K. M. (2015). Lexical and sublexical orthographic processing: an ERP study with skilled and dyslexic adult readers. *Brain and Language*, 141, 16-27. doi:10.1016/j.bandl.2014.11.007
- Assadollahi, R., & Pulvermuller, F. (2001). Neuromagnetic evidence for early access to cognitive representations. *Neuroreport*, 12, 207-213.
- Baillet, S. (2017). Magnetoencephalography for brain electrophysiology and imaging. *Nature Neuroscience*, 20, 327-339. doi:10.1038/nn.4504
- Baker, C. I., Liu, J., Wald, L. L., Kwong, K. K., Benner, T., & Kanwisher, N. (2007). Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 9087-9092. doi:10.1073/pnas.0703300104
- Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental*

Psychology: Human Perception and Performance, 10, 340-357. doi:10.1037/0096-1523.10.3.340

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, 133, 283-316. doi:10.1037/0096-3445.133.2.283

Balota, D. A., & Yap, M. J. (2006). Attentional Control and Flexible Lexical Processing: Explorations of the Magic Moment of Word Recognition. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 229-258). New York, NY: Psychology Press.

Balota, D. A., Yap, M. J., & Cortese, M. J. (2006). Visual word recognition: The journey from features to meaning (a travel update). *Handbook of Psycholinguistics*, 2, 285-375.

Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., . . . Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, 85, 390-401. doi:10.1016/j.neuron.2014.12.018

Ben-Shachar, M., Dougherty, R. F., & Wandell, B. A. (2007). White matter pathways in reading. *Current Opinion in Neurobiology*, 17, 258-270. doi:10.1016/j.conb.2007.03.006

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, 11, 235-260. doi:10.1162/089892999563373

Broca P. (1861). Perte de la Parole, Ramollissement Chronique et Destruction Partielle du Lobe Antérieur Gauche du Cerveau. *Bulletin de la Société Anthropologique*, 2, 235-238.

- Binder, J. R., Medler, D. A., Westbury, C. F., Liebenthal, E., & Buchanan, L. (2006). Tuning of the human left fusiform gyrus to sublexical orthographic structure. *Neuroimage*, 33, 739-748. doi:10.1016/j.neuroimage.2006.06.053
- 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. doi:10.1523/JNEUROSCI.0864-05.2005
- Bitan, T., Burman, D. D., Lu, D., Cone, N. E., Gitelman, D. R., Mesulam, M. M., & Booth, J. R. (2006). Weaker top-down modulation from the left inferior frontal gyrus in children. *Neuroimage*, 33, 991-998. doi:10.1016/j.neuroimage.2006.07.007
- Boets, B., Op de Beeck, H. P., Vandermosten, M., Scott, S. K., Gillebert, C. R., Mantini, D., . . . Ghesquiere, P. (2013). Intact but less accessible phonetic representations in adults with dyslexia. *Science*, 342, 1251-1254. doi:10.1126/science.1244333
- 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. doi:10.1002/hbm.20124
- Boly, M., Garrido, M. I., Gosseries, O., Bruno, M. A., Boveroux, P., Schnakers, C., . . . Friston, K. (2011). Preserved feedforward but impaired top-down processes in the vegetative state. *Science*, 332, 858-862. doi:10.1126/science.1202043
- Booth, J. R., Bebko, G., Burman, D. D., & Bitan, T. (2007). Children with reading disorder show modality independent brain abnormalities during semantic tasks. *Neuropsychologia*, 45, 775-783. doi:10.1016/j.neuropsychologia.2006.08.015
- Booth, J. R., Mehdiratta, N., Burman, D. D., & Bitan, T. (2008). Developmental increases in effective connectivity to brain regions involved in phonological processing during tasks

- with orthographic demands. *Brain Research*, 1189, 78-89.
doi:10.1016/j.brainres.2007.10.080
- Bouhali, F., Thiebaut de Schotten, M., Pinel, P., Poupon, C., Mangin, J. F., Dehaene, S., & Cohen, L. (2014). Anatomical connections of the visual word form area. *Journal of Neuroscience*, 34, 15402-15414. doi:10.1523/JNEUROSCI.4918-13.2014
- Brauer, J., Anwender, A., Perani, D., & Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain and Language*, 127, 289-295.
doi:10.1016/j.bandl.2013.03.001
- Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., & Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Human Brain Mapping*, 30, 1977-1989.
doi:10.1002/hbm.20643
- Cai, Q., Lavidor, M., Brysbaert, M., Paulignan, Y., & Nazir, T. A. (2008). Cerebral lateralization of frontal lobe language processes and lateralization of the posterior visual word processing system. *Journal of Cognitive Neuroscience*, 20(4), 672-681.
doi:10.1162/jocn.2008.20043
- Cao, F., Bitan, T., & Booth, J. R. (2008). Effective brain connectivity in children with reading difficulties during phonological processing. *Brain and Language*, 107(2), 91-101.
doi:10.1016/j.bandl.2007.12.009
- Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, 18, 90-98.
doi:10.1016/j.tics.2013.11.005

- 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. doi:10.1162/jocn.2007.19.3.433
- 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. doi:10.1002/hbm.20236
- Carreiras, M., Seghier, M. L., Baquero, S., Estevez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature*, 461, 983-986. doi:10.1038/nature08461
- Carreiras, M., Vergara, M., & Barber, H. (2005). Early event-related potential effects of syllabic processing during visual word recognition. *Journal of Cognitive Neuroscience*, 17, 1803-1817. doi:10.1162/089892905774589217
- Castles, A., & Friedmann, N. (2014). Developmental Dyslexia and the Phonological Deficit Hypothesis. *Mind and Language*, 29, 270-285. doi:10.1111/mila.12050
- Catani, M., Jones, D. K., & ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, 57, 8-16. doi:10.1002/ana.20319
- Cattinelli, I., Borghese, N. A., Gallucci, M., & Paulesu, E. (2013). Reading the reading brain: A new meta-analysis of functional imaging data on reading. *Journal of Neurolinguistics*, 26, 214-238. doi:10.1016/j.jneuroling.2012.08.001
- Chen, C. C., Kiebel, S. J., & Friston, K. J. (2008). Dynamic causal modelling of induced responses. *Neuroimage*, 41, 1293-1312. doi:10.1016/j.neuroimage.2008.03.026
- Chen, Y., Davis, M. H., Pulvermuller, F., & Hauk, O. (2013). Task modulation of brain responses in visual word recognition as studied using EEG/MEG and fMRI. *Frontiers in Human Neuroscience*, 7. doi:10.3389/fnhum.2013.00376

- Chen, Y., Davis, M. H., Pulvermuller, F., & Hauk, O. (2015). Early visual word processing is flexible: Evidence from spatiotemporal brain dynamics. *Journal of Cognitive Neuroscience*, 27, 1738-1751. doi:10.1162/jocn_a_00815
- Coch, D. (2015). The N400 and the fourth grade shift. *Developmental Science*, 18, 254-269. doi:10.1111/desc.12212
- Coch, D., & Benoit, C. (2015). N400 event-related potential and standardized measures of reading in late elementary school children: Correlated or independent? *Mind, Brain, and Education*, 9, 145-153. doi:10.1111/mbe.12083
- Coch, D., & Meade, G. (2016). N1 and P2 to words and wordlike stimuli in late elementary school children and adults. *Psychophysiology*, 53, 115-128. doi:10.1111/psyp.12567
- Coch, D., & Mitra, P. (2010). Word and pseudoword superiority effects reflected in the ERP waveform. *Brain Research*, 1329, 159-174. doi:10.1016/j.brainres.2010.02.084
- 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, 291-307. doi:10.1093/brain/123.2.291
- 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. doi:10.1016/j.neuroimage.2004.07.052
- 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. doi:10.1037/0033-295X.108.1.204
- 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.
doi:10.1371/journal.pone.0005359
- Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. M. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research*, 1084, 89-103. doi:10.1016/j.brainres.2006.02.010
- Daunizeau, J., David, O., & Stephan, K. E. (2011). Dynamic causal modelling: a critical review of the biophysical and statistical foundations. *Neuroimage*, 58, 312-322. doi:10.1016/j.neuroimage.2009.11.062
- David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., & Friston, K. J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *Neuroimage*, 30, 1255-1272. doi:10.1016/j.neuroimage.2005.10.045
- Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. *Psychophysiology*, 41, 60-74. doi:10.1111/1469-8986.00120
- Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56, 384-398. doi:10.1016/j.neuron.2007.10.004
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15, 254-262. doi:10.1016/j.tics.2011.04.003
- Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences*, 9, 335-341. doi:10.1016/j.tics.2005.05.004
- Dejerine J. (1895). Anatomie des centres nerveux. Vol 1. Paris: Rueff et Cie.

- Dunabeitia, J. A., Dimitropoulou, M., Grainger, J., Hernandez, J. A., & Carreiras, M. (2012). Differential sensitivity of letters, numbers, and symbols to character transpositions. *Journal of Cognitive Neuroscience*, 24, 1610-1624. doi:10.1162/jocn_a_00180
- Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping*, 30, 2907-2926. doi:10.1002/hbm.20718
- Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., . . . Constable, R. T. (2014). Disruption of functional networks in dyslexia: A whole-brain, data-driven analysis of connectivity. *Biological Psychiatry*, 76, 397-404. doi:10.1016/j.biopsych.2013.08.031
- Flowers, D. L., Jones, K., Noble, K., VanMeter, J., Zeffiro, T. A., Wood, F. B., & Eden, G. F. (2004). Attention to single letters activates left extrastriate cortex. *Neuroimage*, 21, 829-839. doi:10.1016/j.neuroimage.2003.10.002
- Friedmann, N., & Coltheart, M. (2016). Types of developmental dyslexia. *Handbook of communication disorders: Theoretical, empirical, and applied linguistics perspectives*.
- Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron*, 88, 220-235. doi:10.1016/j.neuron.2015.09.034
- Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2, 56-78. doi:10.1002/hbm.460020107
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Review Neuroscience*, 11, 127-138. doi:10.1038/nrn2787
- Friston, K. J. (2002). Bayesian estimation of dynamical systems: An application to fMRI. *Neuroimage*, 16, 513-530. doi:10.1006/nimg.2001.1044

- Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity*, *1*, 13-36. doi:10.1089/brain.2011.0008
- Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *Neuroimage*, *19*, 1273-1302. doi:10.1016/S1053-8119(03)00202-7
- Fujimaki, N., Hayakawa, T., Ihara, A., Wei, Q., Munetsuna, S., Terazono, Y., . . . Murata, T. (2009). Early neural activation for lexico-semantic access in the left anterior temporal area analyzed by an fMRI-assisted MEG multidipole method. *Neuroimage*, *44*, 1093-1102. doi:10.1016/j.neuroimage.2008.10.021
- Garrido, M. I., Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2007). Evoked brain responses are generated by feedback loops. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 20961-20966. doi:10.1073/pnas.0706274105
- Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E., & Friston, K. J. (2007). Dynamic causal modelling of evoked potentials: A reproducibility study. *Neuroimage*, *36*, 571-580. doi:10.1016/j.neuroimage.2007.03.014
- Ge, J. Q., Peng, G., Lyu, B. J., Wang, Y., Zhuo, Y., Niu, Z. D., . . . Gao, J. H. (2015). Cross-language differences in the brain network subserving intelligible speech. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 2972-2977. doi:10.1073/pnas.1416000112
- Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*, *15*, 870-878. doi:10.1006/nimg.2001.1037
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain*, *88*, 585-585.

- Grainger, J., & Holcomb, P. J. (2009). Watching the Word Go by: On the time-course of component processes in visual word recognition. *Language and Linguistics Compass*, 3, 128-156. doi:10.1111/j.1749-818x.2008.00121.x
- Grainger, J., & Ziegler, J. C. (2011). A dual-route approach to orthographic processing. *Frontiers in Psychology*, 2, 54. doi:10.3389/fpsyg.2011.00054
- Graves, W. W., Boukrina, O., Mattheiss, S. R., Alexander, E. J., & Baillet, S. (2017). Reversing the standard neural signature of the word-nonword distinction. *Journal of Cognitive Neuroscience*, 29, 79-94. doi:10.1162/jocn_a_01022
- Hall, E. L., Robson, S. E., Morris, P. G., & Brookes, M. J. (2014). The relationship between MEG and fMRI. *Neuroimage*, 102, 80-91. doi:10.1016/j.neuroimage.2014.11.005
- Hampson, M., Tokoglu, F., Sun, Z. D., Schafer, R. J., Skudlarski, P., Gore, J. C., & Constable, R. T. (2006). Connectivity-behavior analysis reveals that functional connectivity between left BA39 and Broca's area varies with reading ability. *Neuroimage*, 31, 513-519. doi:10.1016/j.neuroimage.2005.12.040
- Hancock, R., Richlan, F., & Hoefft, F. (2017). Possible roles for fronto-striatal circuits in reading disorder. *Neuroscience and Biobehavioral Reviews*, 72, 243-260. doi:10.1016/j.neubiorev.2016.10.025
- Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., & Dehaene, S. (2015). Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends in Cognitive Sciences*, 19, 374-382. doi:10.1016/j.tics.2015.05.006
- Hauk, O. (2016). Only time will tell - why temporal information is essential for our neuroscientific understanding of semantics. *Psychonomic Bulletin and Review*, 23, 1072-1079. doi:10.3758/s13423-015-0873-9

- Hauk, O., Coutout, C., Holden, A., & Chen, Y. (2012). The time-course of single-word reading: evidence from fast behavioral and brain responses. *Neuroimage*, 60, 1462-1477. doi:10.1016/j.neuroimage.2012.01.061
- Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage*, 30, 1383-1400. doi:10.1016/j.neuroimage.2005.11.048
- Hauk, O., & Pulvermuller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, 115, 1090-1103. doi:10.1016/j.clinph.2003.12.020
- Heim, S., Eickhoff, S. B., Ischebeck, A. K., Friederici, A. D., Stephan, K. E., & Amunts, K. (2009). Effective connectivity of the left BA 44, BA 45, and inferior temporal gyrus during lexical and phonological decisions identified with DCM. *Human Brain Mapping*, 30, 392-402. doi:10.1002/hbm.20512
- Hoffman, P., Ralph, M. A. L., & Woollams, A. M. (2015). Triangulation of the neurocomputational architecture underpinning reading aloud. *Proceedings of the National Academy of Sciences of the United States of America*, 112, E3719-E3728. doi:10.1073/pnas.1502032112
- Holcomb, P. J., Grainger, J., & O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*, 14, 938-950. doi:10.1162/089892902760191153
- Hon, N., Thompson, R., Sigala, N., & Duncan, J. (2009). Evidence for long-range feedback in target detection: Detection of semantic targets modulates activity in early visual areas. *Neuropsychologia*, 47, 1721-1727. doi:10.1016/j.neuropsychologia.2009.02.011

- 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.
- Houde, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: an fMRI meta-analysis of 52 studies including 842 children. *Developmental Science*, 13, 876-885. doi:10.1111/j.1467-7687.2009.00938.x
- 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. doi:10.1016/j.bandl.2010.10.002
- Jansen, B. H., & Rit, V. G. (1995). Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biological Cybernetics*, 73, 357-366.
- Kado, H., Higuchi, M., Shimogawara, M., Haruta, Y., Adachi, Y., Kawai, J., . . . Uehara, G. (1999). Magnetoencephalogram systems developed at KIT. *IEEE Transactions on Applied Superconductivity*, 9, 4057-4062. doi:10.1109/77.783918
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman brief intelligence test*. John Wiley & Sons, Inc..
- Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavioural Research Methods*, 42, 627-633. doi:10.3758/BRM.42.3.627
- Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: a lexical decision study on 14,000 dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology*, 1, 174. doi:10.3389/fpsyg.2010.00174

- Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavioral Research Methods*, 44, 287-304. doi:10.3758/s13428-011-0118-4
- Kherif, F., Josse, G., & Price, C. J. (2011). Automatic top-down processing explains common left occipito-temporal responses to visual words and objects. *Cerebral Cortex*, 21, 103-114. doi:10.1093/cercor/bhq063
- Kiebel, S. J., David, O., & Friston, K. J. (2006). Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization. *Neuroimage*, 30, 1273-1284. doi:10.1016/j.neuroimage.2005.12.055
- Kiebel, S. J., Garrido, M. I., Moran, R., Chen, C. C., & Friston, K. J. (2009). Dynamic causal modeling for EEG and MEG. *Human Brain Mapping*, 30, 1866-1876. doi:10.1002/hbm.20775
- Kinoshita, S., & Norris, D. (2012). Task-dependent masked priming effects in visual word recognition. *Frontiers in Psychology*, 3, 178. doi:10.3389/fpsyg.2012.00178
- Klein, M., Grainger, J., Wheat, K. L., Millman, R. E., Simpson, M. I., Hansen, P. C., & Cornelissen, P. L. (2015). Early activity in Broca's area during reading reflects fast access to articulatory codes from print. *Cerebral Cortex*, 25, 1715-1723. doi:10.1093/cercor/bht350
- Klingberg, T., Hedehus, M., Temple, E., Salz, T., Gabrieli, J. D. E., Moseley, M. E., & Poldrack, R. A. (2000). Microstructure of temporo-parietal white matter as a basis for reading ability: Evidence from diffusion tensor magnetic resonance imaging. *Neuron*, 25, 493-500. doi:10.1016/S0896-6273(00)80911-3
- Koyama, M. S., Di Martino, A., Zuo, X. N., Kelly, C., Mennes, M., Jutagir, D. R., . . . Milham, M. P. (2011). Resting-state functional connectivity indexes reading competence in

- children and adults. *Journal of Neuroscience*, 31, 8617-8624.
doi:10.1523/JNEUROSCI.4865-10.2011
- Kujala, J., Pammer, K., Cornelissen, P., Roebroek, A., Formisano, E., & Salmelin, R. (2007). Phase coupling in a cerebro-cerebellar network at 8-13 hz during reading. *Cerebral Cortex*, 17, 1476-1485. doi:10.1093/cercor/bhl059
- Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, 62, 621-647. doi:10.1146/annurev.psych.093008.131123
- Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., . . . Fox, P. T. (2005). ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping*, 25, 155-164. doi:10.1002/hbm.20136
- Laszlo, S., & Federmeier, K. D. (2011). The N400 as a snapshot of interactive processing: Evidence from regression analyses of orthographic neighbor and lexical associate effects. *Psychophysiology*, 48, 176-186. doi:10.1111/j.1469-8986.2010.01058.x
- Laszlo, S., & Federmeier, K. D. (2014). Never seem to find the time: Evaluating the physiological time course of visual word recognition with regression analysis of single item ERPs. *Language and Cognitive Processes*, 29, 642-661. doi:10.1080/01690965.2013.866259
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, 9, 920-933. doi:10.1038/nrn2532
- Li, Y., Zhang, L., Xia, Z., Yang, J., Shu, H., & Li, P. (2017). The relationship between intrinsic couplings of the visual word form area with spoken language network and reading

- ability in children and adults. *Frontiers in Human Neuroscience*, 11, 327.
doi:10.3389/fnhum.2017.00327
- Litvak, V., Mattout, J., Kiebel, S., Phillips, C., Henson, R., Kilner, J., . . . Friston, K. (2011). EEG and MEG data analysis in SPM8. *Computational Intelligence and Neuroscience*, 2011, 852961. doi:10.1155/2011/852961
- Liu, C., Zhang, W. T., Tang, Y. Y., Mai, X. Q., Chen, H. C., Tardif, T., & Luo, Y. J. (2008). The visual word form area: Evidence from an fMRI study of implicit processing of Chinese characters. *Neuroimage*, 40, 1350-1361.
doi:10.1016/j.neuroimage.2007.10.014
- Liu, J., Li, J., Rieth, C. A., Huber, D. E., Tian, J., & Lee, K. (2011). A dynamic causal modeling analysis of the effective connectivities underlying top-down letter processing. *Neuropsychologia*, 49, 1177-1186. doi:10.1016/j.neuropsychologia.2011.01.011
- Liu, J., Li, J., Zhang, H., Rieth, C. A., Huber, D. E., Li, W., . . . Tian, J. (2010). Neural correlates of top-down letter processing. *Neuropsychologia*, 48, 636-641.
doi:10.1016/j.neuropsychologia.2009.10.024
- Liu, L., Vira, A., Friedman, E., Minas, J., Bolger, D., Bitan, T., & Booth, J. (2010). Children with reading disability show brain differences in effective connectivity for visual, but not auditory word comprehension. *Plos One*, 5, e13492.
doi:10.1371/journal.pone.0013492
- 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.
doi:10.1038/35084005

- Mahe, G., Zesiger, P., & Laganaro, M. (2015). Beyond the initial 140 ms, lexical decision and reading aloud are different tasks: An ERP study with topographic analysis. *Neuroimage*, 122, 65-72. doi:10.1016/j.neuroimage.2015.07.080
- Maisog, J. M., Einbinder, E. R., Flowers, D. L., Turkeltaub, P. E., & Eden, G. F. (2008). A meta-analysis of functional neuroimaging studies of dyslexia. *Annals of the New York Academy of Sciences*, 1145, 237-259. doi:10.1196/annals.1416.024
- Mani, J., Diehl, B., Piao, Z., Schuele, S. S., Lapresto, E., Liu, P., . . . Luders, H. O. (2008). Evidence for a basal temporal visual language center: cortical stimulation producing pure alexia. *Neurology*, 71, 1621-1627. doi:10.1212/01.wnl.0000334755.32850.f0
- Mano, Q. R., Humphries, C., Desai, R. H., Seidenberg, M. S., Osmon, D. C., Stengel, B. C., & Binder, J. R. (2013). The role of left occipitotemporal cortex in reading: reconciling stimulus, task, and lexicality effects. *Cerebral Cortex*, 23, 988-1001. doi:10.1093/cercor/bhs093
- Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron*, 38, 487-497. doi:10.1016/S0896-6273(03)00197-1
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164, 177-190. doi:10.1016/j.jneumeth.2007.03.024
- Martin, A., Schurz, M., Kronbichler, M., & Richlan, F. (2015). Reading in the brain of children and adults: a meta-analysis of 40 functional magnetic resonance imaging studies. *Human Brain Mapping*, 36, 1963-1981. doi:10.1002/hbm.22749
- Martin, C. D., Nazir, T., Thierry, G., Paulignan, Y., & Demonet, J. F. (2006). Perceptual and lexical effects in letter identification: an event-related potential study of the word superiority effect. *Brain Research*, 1098, 153-160. doi:10.1016/j.brainres.2006.04.097

- Maurer, U., Brandeis, D., & McCandliss, B. D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behavioral and Brain Functions, 1*, 13. doi:10.1186/1744-9081-1-13
- 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. doi:10.1016/S1364-6613(03)00134-7
- McNorgan, C., Chabal, S., O'Young, D., Lukic, S., & Booth, J. R. (2015). Task dependent lexicality effects support interactive models of reading: A meta-analytic neuroimaging review. *Neuropsychologia, 67*, 148-158. doi:10.1016/j.neuropsychologia.2014.12.014
- Melchert, H. C. (2004). *A dictionary of the Lycian language*. Beech Stave Press Inc..
- Michalareas, G., Vezoli, J., van Pelt, S., Schoffelen, J. M., Kennedy, H., & Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron, 89*, 384-397. doi:10.1016/j.neuron.2015.12.018
- Monsell, S., Doyle, M. C., & Haggard, P. N. (1989). Effects of frequency on visual word recognition tasks - Where are they. *Journal of Experimental Psychology: General, 118*, 43-71. doi:10.1037/0096-3445.118.1.43
- Morken, F., Helland, T., Hugdahl, K., & Specht, K. (2017). Reading in dyslexia across literacy development: A longitudinal study of effective connectivity. *Neuroimage, 144*, 92-100. doi:10.1016/j.neuroimage.2016.09.060
- Norris, D. (2013). Models of visual word recognition. *Trends in Cognitive Sciences, 17*, 517-524. doi:10.1016/j.tics.2013.08.003
- Norris, D., & Kinoshita, S. (2008). Perception as evidence accumulation and Bayesian inference: insights from masked priming. *Journal of Experimental Psychology: General, 137*, 434-455. doi:10.1037/a0012799

- Oberhuber, M., Parker, Jones, Hope, T. M., Prejawa, S., Seghier, M. L., Green, D. W., & Price, C. J. (2013). Functionally distinct contributions of the anterior and posterior putamen during sublexical and lexical reading. *Frontiers in Human Neuroscience*, 7, 787. doi:10.3389/fnhum.2013.00787
- Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 9868-9872.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97-113.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011, 156869. doi:10.1155/2011/156869
- Pammer, K. (2009). What can MEG neuroimaging tell us about reading? *Journal of Neurolinguistics*, 22, 266-280. doi:10.1016/j.jneuroling.2008.12.004
- Pammer, K., Hansen, P. C., Kringelbach, M. L., Holliday, I., Barnes, G., Hillebrand, A., . . . Cornelissen, P. L. (2004). Visual word recognition: the first half second. *Neuroimage*, 22(4) 1819-1825. doi:10.1016/j.neuroimage.2004.05.004
- Penny, W. D., Litvak, V., Fuentemilla, L., Duzel, E., & Friston, K. (2009). Dynamic Causal Models for phase coupling. *Journal of Neuroscience Methods*, 183, 19-30. doi:10.1016/j.jneumeth.2009.06.029
- Penny, W. D., Stephan, K. E., Daunizeau, J., Rosa, M. J., Friston, K. J., Schofield, T. M., & Leff, A. P. (2010). Comparing families of dynamic causal models. *PLoS Computational Biology*, 6, e1000709. doi:10.1371/journal.pcbi.1000709

- Penny, W. D., Stephan, K. E., Mechelli, A., & Friston, K. J. (2004). Comparing dynamic causal models. *Neuroimage*, 22, 1157-1172. doi:10.1016/j.neuroimage.2004.03.026
- Perani, D., Saccuman, M. C., Scifo, P., Awander, A., Spada, D., Baldoli, C., . . . Friederici, A. D. (2011). Neural language networks at birth. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 16056-16061. doi:10.1073/pnas.1102991108
- Perrone-Bertolotti, M., Kauffmann, L., Pichat, C., Vidal, J. R., & Baciú, M. (2017). Effective connectivity between ventral occipito-temporal and ventral inferior frontal cortex during lexico-semantic processing: A dynamic causal modeling study. *Frontiers in Human Neuroscience*, 11, 325. doi:10.3389/fnhum.2017.00325
- 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.
- Poch, C., Garrido, M. I., Igoa, J. M., Belinchon, M., Garcia-Morales, I., & Campo, P. (2015). Time-varying effective connectivity during visual object naming as a function of semantic demands. *Journal of Neuroscience*, 35, 8768-8776. doi:10.1523/JNEUROSCI.4888-14.2015
- Price, C. J. (2000). The anatomy of language: contributions from functional neuroimaging. *Journal of Anatomy*, 197, 335-359. doi: Doi 10.1046/J.1469-7580.2000.19730335.X
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, 62, 816-847. doi:10.1016/j.neuroimage.2012.04.062
- Price, C. J., & Devlin, J. T. (2003). The myth of the visual word form area. *Neuroimage*, 19, 473-481. doi: 10.1016/S1053-8119(03)00084-3

- Price, C. J., & Devlin, J. T. (2011). The interactive account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, 15, 246-253. doi:10.1016/j.tics.2011.04.001
- Price, C. J., & Mechelli, A. (2005). Reading and reading disturbance. *Current Opinion in Neurobiology*, 15, 231-238. doi:10.1016/j.conb.2005.03.003
- Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., . . . Shaywitz, B. A. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Reviews*, 6, 207–213. doi:10.1002/1098-2779(2000)6:3<207::AID-MRDD8>3.0.CO;2-P
- Pugh, K. R., Mencl, W. E., Shaywitz, B. A., Shaywitz, S. E., Fulbright, R. K., Constable, R. T., . . . Gore, J. C. (2000). The angular gyrus in developmental dyslexia: task-specific differences in functional connectivity within posterior cortex. *Psychological Science*, 11, 51-56. doi:10.1111/1467-9280.00214
- Pylkkanen, L., & Marantz, A. (2003). Tracking the time course of word recognition with MEG. *Trends in Cognitive Sciences*, 7, 187-189. doi:10.1016/S1364-6613(03)00092-5
- Pylkkanen, L., & McElree, B. (2007). An MEG study of silent meaning. *Journal of Cognitive Neuroscience*, 19, 1905-1921. doi:10.1162/jocn.2007.19.11.1905
- Quinn, A (2014) *Neural dynamics in brain networks during the resting state and visual word recognition*. PhD thesis, University of York.
- Richardson, F. M., Seghier, M. L., Leff, A. P., Thomas, M. S., & Price, C. J. (2011). Multiple routes from occipital to temporal cortices during reading. *Journal of Neuroscience*, 31, 8239-8247. doi:10.1523/JNEUROSCI.6519-10.2011

- Richlan, F., Kronbichler, M., & Wimmer, H. (2009). Functional abnormalities in the dyslexic brain: a quantitative meta-analysis of neuroimaging studies. *Human Brain Mapping*, 30, 3299-3308. doi:10.1002/hbm.20752
- Richlan, F., Kronbichler, M., & Wimmer, H. (2011). Meta-analyzing brain dysfunctions in dyslexic children and adults. *Neuroimage*, 56, 1735-1742. doi:10.1016/j.neuroimage.2011.02.040
- Robinson, A. (2009). *Writing and script: A very short introduction (Vol. 208)*. Oxford University Press.
- Saygin, Z. M., Norton, E. S., Osher, D. E., Beach, S. D., Cyr, A. B., Ozernov-Palchik, O., . . . Gabrieli, J. D. E. (2013). Tracking the roots of reading ability: White matter volume and integrity correlate with phonological awareness in prereading and early-reading kindergarten children. *Journal of Neuroscience*, 33, 13251-13258. doi:10.1523/JNEUROSCI.4383-12.2013
- Schoffelen, J. M., Hulten, A., Lam, N., Marquand, A. F., Udden, J., & Hagoort, P. (2017). Frequency-specific directed interactions in the human brain network for language. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 8083-8088. doi:10.1073/pnas.1703155114
- Schurz, M., Wimmer, H., Richlan, F., Ludersdorfer, P., Klackl, J., & Kronbichler, M. (2015). Resting-state and task-Based functional brain connectivity in developmental dyslexia. *Cerebral Cortex*, 25, 3502-3514. doi:10.1093/cercor/bhu184
- Seghier, M. L., Josse, G., Leff, A. P., & Price, C. J. (2011). Lateralization is predicted by reduced coupling from the left to right prefrontal cortex during semantic decisions on written words. *Cerebral Cortex*, 21, 1519-1531. doi:10.1093/cercor/bhq203

- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568. doi:10.1037/0033-295X.96.4.523
- Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: eye movements and event-related potentials. *Trends in Cognitive Sciences*, 7, 489-493. doi:10.1016/j.tics.2003.09.010
- Sereno, S. C., Rayner, K., & Posner, M. I. (1998). Establishing a time-line of word recognition: evidence from eye movements and event-related potentials. *Neuroreport*, 9, 2195-2200.
- Shaywitz, B. A., Shaywitz, S. E., Pugh, K. R., Mencl, W. E., Fulbright, R. K., Skudlarski, P., . . . Gore, J. C. (2002). Disruption of posterior brain systems for reading in children with developmental dyslexia. *Biological Psychiatry*, 52, 101-110. doi:10.1016/S0006-3223(02)01365-3
- Skrandies, W. (1990). Global field power and topographic similarity. *Brain Topography*, 3, 137-141. doi:10.1007/BF01128870
- Spironelli, C., & Angrilli, A. (2007). Influence of phonological, semantic and orthographic tasks on the early linguistic components N150 and N350. *International Journal of Psychophysiology*, 64, 190-198. doi:10.1016/j.ijpsycho.2007.02.002
- Starrfelt, R., & Gerlach, C. (2007). The visual what for area: words and pictures in the left fusiform gyrus. *Neuroimage*, 35, 334-342. doi:10.1016/j.neuroimage.2006.12.003
- Strijkers, K., Bertrand, D., & Grainger, J. (2015). Seeing the same words differently: the time course of automaticity and top-down intention in reading. *Journal of Cognitive Neuroscience*, 27, 1542-1551. doi:10.1162/jocn_a_00797

- Strijkers, K., Yum, Y. N., Grainger, J., & Holcomb, P. J. (2011). Early goal-directed top-down influences in the production of speech. *Frontiers in Psychology*, 2, 371. doi:10.3389/fpsyg.2011.00371
- Taft, M., & van Graan, F. (1998). Lack of phonological mediation in a semantic categorization task. *Journal of Memory and Language*, 38, 203-224. doi:10.1006/jmla.1997.2538
- 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, 2119-2132. doi:10.1093/brain/122.11.2119
- Taylor, J. S., Rastle, K., & Davis, M. H. (2013). Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. *Psychological Bulletin*, 139, 766-791. doi:10.1037/a0030266
- Thesen, T., McDonald, C. R., Carlson, C., Doyle, W., Cash, S., Sherfey, J., . . . Halgren, E. (2012). Sequential then interactive processing of letters and words in the left fusiform gyrus. *Nature Communications*, 3, 1284. doi:10.1038/ncomms2220
- Thiebaut de Schotten, M., Cohen, L., Amemiya, E., Braga, L. W., & Dehaene, S. (2014). Learning to read improves the structure of the arcuate fasciculus. *Cerebral Cortex*, 24, 989-995. doi:10.1093/cercor/bhs383
- Torgesen, J. K., Rashotte, C. A., & Wagner, R. K. (1999). *TOWRE: Test of word reading efficiency*. Austin, TX: Pro-ed.
- Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *Neuroimage*, 16, 765-780. doi:10.1006/nimg.2002.1131

- Twomey, T., Kawabata Duncan, K. J., Price, C. J., & Devlin, J. T. (2011). Top-down modulation of ventral occipito-temporal responses during visual word recognition. *Neuroimage*, 55, 1242-1251. doi:10.1016/j.neuroimage.2011.01.001
- Uehara, G., Adachi, Y., Kawai, J., Shimogawara, M., Higuchi, M., Haruta, Y., . . . Kado, H. (2003). Multi-channel SQUID systems for biomagnetic measurement. *IEICE Transactions on Electronics*, E86c, 43-54.
- Valdois, S., Carbonnel, S., Juphard, A., Baciú, M., Ans, B., Peyrin, C., & Segebarth, C. (2006). Polysyllabic pseudo-word processing in reading and lexical decision: converging evidence from behavioral data, connectionist simulations and functional MRI. *Brain Research*, 1085, 149-162. doi:10.1016/j.brainres.2006.02.049
- Van der Haegen, L., Cai, Q., & Brysbaert, M. (2012). Colateralization of Broca's area and the visual word form area in left-handers: fMRI evidence. *Brain and Language*, 122, 171-178. doi:10.1016/j.bandl.2011.11.004
- van der Mark, S., Klaver, P., Bucher, K., Maurer, U., Schulz, E., Brem, S., . . . 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. doi:10.1016/j.neuroimage.2010.10.002
- Vandermosten, M., Boets, B., Wouters, J., & Ghesquiere, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neuroscience and Biobehavioral Reviews*, 36, 1532-1552. doi:10.1016/j.neubiorev.2012.04.002
- Vartiainen, J., Liljestrom, M., Koskinen, M., Renvall, H., & Salmelin, R. (2011). Functional magnetic resonance imaging blood oxygenation level-dependent signal and magnetoencephalography evoked responses yield different neural functionality in

- reading. *Journal of Neuroscience*, 31, 1048-1058. doi:10.1523/JNEUROSCI.3113-10.2011
- Vartiainen, J., Parviainen, T., & Salmelin, R. (2009). Spatiotemporal convergence of semantic processing in reading and speech perception. *Journal of Neuroscience*, 29, 9271-9280. doi:10.1523/JNEUROSCI.5860-08.2009
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45, 2-40. doi:10.1046/j.0021-9630.2003.00305.x
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. *Neuron*, 55, 143-156. doi:10.1016/j.neuron.2007.05.031
- Vogel, A. C., Miezin, F. M., Petersen, S. E., & Schlaggar, B. L. (2012). The putative visual word form area is functionally connected to the dorsal attention network. *Cerebral Cortex*, 22, 537-549. doi:10.1093/cercor/bhr100
- Wager, T. D., Keller, M. C., Lacey, S. C., & Jonides, J. (2005). Increased sensitivity in neuroimaging analyses using robust regression. *Neuroimage*, 26, 99-113. doi:10.1016/j.neuroimage.2005.01.011
- Wandell, B. A., Rauschecker, A. M., & Yeatman, J. D. (2012). Learning to see words. *Annuals of Review Psychology*, 63, 31-53. doi:10.1146/annurev-psych-120710-100434
- Wang, F., & Maurer, U. (2017). Top-down modulation of early print-tuned neural activity in reading. *Neuropsychologia*, 102, 29-38. doi:10.1016/j.neuropsychologia.2017.05.028

- Wang, X., Han, Z., He, Y., Liu, L., & Bi, Y. (2012). Resting-state functional connectivity patterns predict Chinese word reading competency. *PLoS One*, 7, e44848. doi:10.1371/journal.pone.0044848
- Wang, X., Yang, J., Yang, J., Mencl, W. E., Shu, H., & Zevin, J. D. (2015). Language differences in the brain network for reading in naturalistic story reading and lexical decision. *PLoS One*, 10, e0124388. doi:10.1371/journal.pone.0124388
- Wernicke C. (1874). Der aphasische Symptomencomplex. *Breslau: Cohen and Weigert*.
- Wheat, K. L., Cornelissen, P. L., Frost, S. J., & Hansen, P. C. (2010). During visual word recognition, phonology is accessed within 100 ms and may be mediated by a speech production code: evidence from magnetoencephalography. *Journal of Neuroscience*, 30, 5229-5233. doi:10.1523/JNEUROSCI.4448-09.2010
- Woodhead, Z. V., Barnes, G. R., Penny, W., Moran, R., Teki, S., Price, C. J., & Leff, A. P. (2014). Reading front to back: MEG evidence for early feedback effects during word recognition. *Cerebral Cortex*, 24, 817-825. doi:10.1093/cercor/bhs365
- Wright, N. D., Mechelli, A., Noppeney, U., Veltman, D. J., Rombouts, S. A., Glensman, J., . . . Price, C. J. (2008). Selective activation around the left occipito-temporal sulcus for words relative to pictures: individual variability or false positives? *Human Brain Mapping*, 29, 986-1000. doi:10.1002/hbm.20443
- Xu, M., Wang, T., Chen, S., Fox, P. T., & Tan, L. H. (2015). Effective connectivity of brain regions related to visual word recognition: An fMRI study of Chinese reading. *Human Brain Mapping*, 36, 2580-2591. doi:10.1002/hbm.22792
- Yang, J., Wang, X., Shu, H., & Zevin, J. D. (2012). Task by stimulus interactions in brain responses during Chinese character processing. *Neuroimage*, 60, 979-990. doi:10.1016/j.neuroimage.2012.01.036

- Yang, J., & Zevin, J. (2014). The impact of task demand on visual word recognition. *Neuroscience*, 272, 102-115. doi:10.1016/j.neuroscience.2014.04.044
- Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, 60, 502-529. doi:10.1016/j.jml.2009.02.001
- Yeatman, J. D., Dougherty, R. F., Ben-Shachar, M., & Wandell, B. A. (2012). Development of white matter and reading skills. *Proceedings of the National Academy of Sciences of the United States of America*, 109, E3045-E3053. doi:10.1073/pnas.1206792109
- Yeatman, J. D., Dougherty, R. F., Rykhlevskaia, E., Sherbondy, A. J., Deutsch, G. K., Wandell, B. A., & Ben-Shachar, M. (2011). Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. *Journal of Cognitive Neuroscience*, 23, 3304-3317. doi:10.1162/jocn_a_00061
- 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. doi:10.1016/j.bandl.2012.04.010
- Yum, Y. N., Law, S. P., Su, I. F., Lau, K. Y., & Mo, K. N. (2014). An ERP study of effects of regularity and consistency in delayed naming and lexicality judgment in a logographic writing system. *Frontiers in Psychology*, 5, 315. doi:10.3389/fpsyg.2014.00315
- Yvert, G., Perrone-Bertolotti, M., Baciou, M., & David, O. (2012). Dynamic causal modeling of spatiotemporal integration of phonological and semantic processes: an electroencephalographic study. *Journal of Neuroscience*, 32, 4297-4306. doi:10.1523/JNEUROSCI.6434-11.2012
- Zhao, R., Fan, R., Liu, M. X., Wang, X. J., & Yang, J. F. (2017). Rethinking the function of brain regions for reading Chinese characters in a meta-analysis of fMRI studies. *Journal of Neurolinguistics*, 44, 120-133. doi:10.1016/j.jneuroling.2017.04.001

- Zhu, L., Nie, Y., Chang, C., Gao, J. H., & Niu, Z. (2014). Different patterns and development characteristics of processing written logographic characters and alphabetic words: an ALE meta-analysis. *Human Brain Mapping*, 35, 2607-2618. doi:10.1002/hbm.22354
- 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. doi:10.1037/0033-2909.131.1.3

Ethics Approval

Office of the Deputy Vice-Chancellor
(Research)

Research Office
Research Hub, Building C5C East
Macquarie University
NSW 2109 Australia
T: +61 (2) 9850 4459
<http://www.research.mq.edu.au/>
ABN 90 952 801 237



MACQUARIE
University
SYDNEY · AUSTRALIA

28 April 2015

Professor Anne Castles
Department of Cognitive Science
Faculty of Human Sciences
Macquarie University
NSW 2109

Dear Professor Castles

Reference No: 5201500186

Title: *Top-down modulation of interior frontal gyrus to ventral occipito-temporal cortex during the early stages of visual word processing*

Thank you for submitting the above application for ethical and scientific review. Your application was considered by the Macquarie University Human Research Ethics Committee (HREC (Medical Sciences)) at its meeting on 26 March 2015 at which further information was requested to be reviewed by the Ethics Secretariat.

The requested information was received with correspondence on 20 April and 22 April 2015.

I am pleased to advise that ethical and scientific approval has been granted for this project to be conducted at:

- Macquarie University

This research meets the requirements set out in the *National Statement on Ethical Conduct in Human Research* (2007 – Updated March 2014) (the *National Statement*).

This letter constitutes ethical and scientific approval only.

Standard Conditions of Approval:

1. Continuing compliance with the requirements of the *National Statement*, which is available at the following website:

<http://www.nhmrc.gov.au/book/national-statement-ethical-conduct-human-research>

2. This approval is valid for five (5) years, subject to the submission of annual reports. Please submit your reports on the anniversary of the approval for this protocol.

3. All adverse events, including events which might affect the continued ethical and scientific acceptability of the project, must be reported to the HREC within 72 hours.

4. Proposed changes to the protocol must be submitted to the Committee for approval before implementation.

It is the responsibility of the Chief investigator to retain a copy of all documentation related to this project and to forward a copy of this approval letter to all personnel listed on the project.

Should you have any queries regarding your project, please contact the Ethics Secretariat on 9850 4194 or by email ethics.secretariat@mq.edu.au

The HREC (Medical Sciences) Terms of Reference and Standard Operating Procedures are available from the Research Office website at:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics

The HREC (Medical Sciences) wishes you every success in your research.

Yours sincerely



Professor Tony Eyers

Chair, Macquarie University Human Research Ethics Committee (Medical Sciences)

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice*.

cc. Mr Yu Li

Details of this approval are as follows:

Approval Date: 22 April 2015

The following documentation has been reviewed and approved by the HREC (Medical Sciences):

Documents reviewed	Version no.	Date
Macquarie University Ethics Application Form	2.3	July 2013
Correspondence from Mr Yu Li responding to the issues raised by the HREC (Medical Sciences)		Received 20/4/2015 & 22/4/2015
MQ Participant Information and Consent Form (PICF)	2	22/4/2015
Macquarie Medical Imaging MRI Neuroimaging Research Request Form		
Language Background Questionnaire	2	20/4/2015
Participant Stimuli – Real words, Pseudowords, Consonant strings	2	20/4/2015
Edinburgh Handedness Inventory		
MEG Experiment Procedure	2	20/4/2015