Mapping dynamic brain connectivity using EEG, TMS, and Transfer Entropy

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2016

Chris Repper-Day

School of Biological Sciences
List of Figures ................................................................................ 6
List of Tables .................................................................................. 8
List of Abbreviations ...................................................................... 9
Abstract ....................................................................................... 11
Declaration .................................................................................. 12
Copyright Statement .................................................................... 13
Contributions of Co-Authors ......................................................... 14
Acknowledgments ........................................................................ 15
Thesis Overview ........................................................................... 16
Chapter 1 – General Introduction ................................................ 17
  1.1 Brain-mapping techniques ........................................................ 19
    1.1.1 Haemodynamic and metabolic methods .............................................. 20
    1.1.2 Electrophysiological methods ............................................................. 21
    1.1.3 Transcranial Magnetic Stimulation ...................................................... 22
    1.1.4 Combining brain mapping techniques ................................................... 23
  1.2 Quantification of functional and effective connectivity .................... 24
    1.2.1 Correlation/Covariance ................................................................. 24
    1.2.2 Coherence ..................................................................................... 25
    1.2.3 Phase synchronisation .................................................................... 25
    1.2.4 Mutual Information ....................................................................... 26
    1.2.5 Dynamic Causal Modelling .............................................................. 27
    1.2.6 Granger causality measures .............................................................. 28
  1.3 Information Theory and Transfer Entropy ...................................... 29
    1.3.1 Information Theory ...................................................................... 29
    1.3.2 Transfer Entropy ........................................................................ 31
    1.3.3 Investigating effective connectivity within EEG data using Transfer Entropy .. 32
  1.4 EEG source reconstruction ........................................................ 34
    1.4.1 Forward Models ........................................................................ 35
1.4.2 Inverse Solutions ........................................................................................................ 36

1.5 Semantic and number processing ..................................................................... 39
  1.5.1 Semantic cognition in the brain ..................................................................... 39
  1.5.2 Number processing in the brain ..................................................................... 41

1.6 Aims of the thesis ........................................................................................................ 42

Chapter 2 - ERP analyses of semantic category and number magnitude judgement tasks ......................................................................................................................... 43

2.1 Abstract ................................................................................................................ 44

2.2 Introduction .......................................................................................................... 45
  2.2.1 Aims of the current study ............................................................................. 46

2.3 Methods ................................................................................................................ 47
  2.3.1 Participants .................................................................................................. 47
  2.3.2 Stimuli ......................................................................................................... 47
  2.3.3 Experimental Paradigm .............................................................................. 48
  2.3.4 EEG Data Acquisition .............................................................................. 49
  2.3.5 Data Pre-processing .................................................................................. 50
  2.3.6 Data Analysis ........................................................................................... 50

2.4 Results ................................................................................................................ 51
  2.4.1 Behavioural Results .................................................................................. 51
  2.4.2 Cluster Analysis Results ........................................................................... 52
  2.4.3 Mean amplitude analysis of N400 and N365 ............................................. 54
  2.4.4 Source Analysis ....................................................................................... 58

2.5 Discussion ............................................................................................................. 59
  2.5.1 Semantic category judgment ..................................................................... 60
  2.5.2 Number magnitude judgment .................................................................... 61
  2.5.3 Conclusion ............................................................................................... 62

Chapter 3 - Trial-based transfer entropy analysis of semantic category and number magnitude judgement tasks ................................................................. 63

3.1 Abstract ................................................................................................................. 64

3.2 Introduction .......................................................................................................... 65
List of Figures

Figure 1.1. The principle of TE causality analysis .......................................................... 32
Figure 2.1. Word and number trial sequences .................................................................. 49
Figure 2.2. T-value scalp topographies of significant clusters and average ERP after targets .................................................................................................................. 53
Figure 2.3. Average ERP and mean amplitude of the N400 at all stimuli ....................... 55
Figure 2.4. Average ERP and mean amplitude of the N365 at all stimuli ....................... 56
Figure 2.5. Average ERP from left and right parietal-occipital electrode clusters .......... 57
Figure 2.6. Source reconstructions of N365 and N400 components ............................... 58
Figure 3.1. The principle of TE causality analysis .......................................................... 70
Figure 3.2. Word and number trial sequences .................................................................. 73
Figure 3.3. BEM model and subject source grid ............................................................ 75
Figure 3.4. Direct discretisation method ......................................................................... 77
Figure 3.5. Trial-based window TE analysis between sources ....................................... 79
Figure 3.6. Grand average TE Matrices within 200ms windows following word targets ... 82
Figure 3.7. Grand average TE Matrices within 200ms windows following number targets .................................................................................................................. 83
Figure 3.8. Information exchange at each time-window ............................................... 84
Figure 3.9. Intra-regional TE at each time-window ....................................................... 85
Figure 3.10. Regional Source Signal Power at each time window ................................ 88
Figure 3.11. Information flow between temporal sources and the rest of the brain ....... 90
Figure 3.12. Information flow between parietal sources and the rest of the brain .......... 91
Figure 4.1. Sites of both temporal and parietal TMS ..................................................... 106
Figure 4.2. TMS-EEG data cleaning pipeline ............................................................... 108
Figure 4.3. Trial-based window TE analysis between sources ...................................... 112
Figure 4.4. Grand average of TMS-evoked activity ...................................................... 113
Figure 4.5. Grand average source TE matrices pre- and post-TMS to lATL ................. 115
Figure 4.6. Grand average source TE matrices pre- and post-TMS to rATL ................. 116
Figure 4.7. Grand average source TE matrices pre- and post-TMS to lIPS ................. 117
Figure 4.8. Grand average source TE matrices pre- and post-TMS to rIPS .......... 118
Figure 4.9. Pre- and post-TMS mean source TE within windows of interest .......... 120
Figure 4.10. Distribution of baseline sending source TE from IATL TMS-proxy source to rest of the brain ................................................................. 121
Figure 4.11. Information sent from left and right ATL TMS-proxy sources at 0-200ms post-TMS ......................................................................................... 124
Figure 4.12. Information sent from left and right ATL TMS-proxy sources at 200-400ms post-TMS ......................................................................................... 125
Figure 4.13. Information sent from left and right ATL TMS-proxy sources at 400-600ms post-TMS ......................................................................................... 126
Figure 4.14. Information sent from left and right IPS TMS-proxy sources at 0-200ms post-TMS ......................................................................................... 128
Figure 4.15. Information sent from left and right IPS TMS-proxy sources at 200-400ms post-TMS ......................................................................................... 129
Figure 4.16. Information sent from left and right IPS TMS-proxy sources at 400-600ms post-TMS ......................................................................................... 130
List of Tables

Table 3.1. Mean intra-regional source TE following both words and numbers in all 3 windows............................................................................................................86

Table 4.1. Mean source TE between windows of interest after TMS to all 4 sites........ 119

Table 4.2. Descriptive statistics of pre-TMS sending source TE at all 4 TMS-proxies.... 121
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAL</td>
<td>Automated Atlas Labelling</td>
</tr>
<tr>
<td>AIR</td>
<td>Average Information Received</td>
</tr>
<tr>
<td>AIS</td>
<td>Average Information Sent</td>
</tr>
<tr>
<td>ATL</td>
<td>Anterior Temporal Lobe</td>
</tr>
<tr>
<td>BEM</td>
<td>Boundary Element Method</td>
</tr>
<tr>
<td>BOLD</td>
<td>Blood-oxygen-level dependent</td>
</tr>
<tr>
<td>DCM</td>
<td>Dynamic Causal Modelling</td>
</tr>
<tr>
<td>DICS</td>
<td>Dynamic Imaging of Coherent Sources</td>
</tr>
<tr>
<td>DLPFC</td>
<td>Dorsolateral Prefrontal Cortex</td>
</tr>
<tr>
<td>DTI</td>
<td>Diffusion Tensor Imaging</td>
</tr>
<tr>
<td>dsTMS</td>
<td>Dual-site TMS</td>
</tr>
<tr>
<td>EC</td>
<td>Effective Connectivity</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalography</td>
</tr>
<tr>
<td>ERF</td>
<td>Event-related field</td>
</tr>
<tr>
<td>ERP</td>
<td>Event-related Potential</td>
</tr>
<tr>
<td>FC</td>
<td>Functional Connectivity</td>
</tr>
<tr>
<td>FDR</td>
<td>False Discovery Rate</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>GC</td>
<td>Granger Causality</td>
</tr>
<tr>
<td>GM</td>
<td>Grey Matter</td>
</tr>
<tr>
<td>ICA</td>
<td>Independent Component Analysis</td>
</tr>
<tr>
<td>IPL</td>
<td>Inferior Parietal Lobule</td>
</tr>
<tr>
<td>IPS</td>
<td>Intra-parietal Sulcus</td>
</tr>
<tr>
<td>LCMV</td>
<td>Linearly Constrained Minimum Variance</td>
</tr>
<tr>
<td>LORETA</td>
<td>Low Resolution Electromagnetic Tomography</td>
</tr>
<tr>
<td>LP</td>
<td>Left Parietal</td>
</tr>
<tr>
<td>LT</td>
<td>Left Temporal</td>
</tr>
<tr>
<td>M1</td>
<td>Primary Motor Cortex</td>
</tr>
<tr>
<td>M1\text{HAND}</td>
<td>Hand region of primary motor cortex</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>MEP</td>
<td>Motor-evoked Potential</td>
</tr>
<tr>
<td>MI</td>
<td>Mutual Information</td>
</tr>
<tr>
<td>MNE</td>
<td>Minimum Norm Estimation</td>
</tr>
<tr>
<td>MNI</td>
<td>Montreal Neurological Institute</td>
</tr>
<tr>
<td>MT</td>
<td>Motor Threshold</td>
</tr>
<tr>
<td>MVAR</td>
<td>Multivariate Autoregressive Model</td>
</tr>
<tr>
<td>PLV</td>
<td>Phase Locking Value</td>
</tr>
<tr>
<td>PS</td>
<td>Phase Synchronisation</td>
</tr>
<tr>
<td>RP</td>
<td>Right Parietal</td>
</tr>
<tr>
<td>RT</td>
<td>Right Temporal</td>
</tr>
<tr>
<td>rTMS</td>
<td>Repetitive TMS</td>
</tr>
<tr>
<td>SLF</td>
<td>Superior Longitudinal Fasciculus</td>
</tr>
<tr>
<td>TBS</td>
<td>Theta Burst Stimulation</td>
</tr>
<tr>
<td>TE</td>
<td>Transfer Entropy</td>
</tr>
<tr>
<td>TEP</td>
<td>TMS-evoked Potential</td>
</tr>
<tr>
<td>TMS</td>
<td>Transcranial Magnetic Stimulation</td>
</tr>
<tr>
<td>WM</td>
<td>Working Memory</td>
</tr>
</tbody>
</table>
Abstract

Mapping dynamic brain connectivity using EEG, TMS, and Transfer Entropy

Chris Repper-Day

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

To understand how the brain functions, we must investigate the transient interactions that underpin communication between cortical regions. EEG possesses the optimal temporal resolution to capture functional connectivity, but it lacks the spatial resolution to identify the cortical locations responsible. To circumvent this problem electrophysiological connectivity should be investigated at the source level. There are many quantifiers of connectivity applied to EEG data, but some are not sensitive to the direct, or indirect, influence of one region over another, and others require the specification of a priori models so are unsuitable for exploratory analyses. Transfer Entropy (TE) can be used to infer the direction of linear and non-linear information exchange between signals over a range of time-delays within EEG data. This thesis explores the creation of a new method of mapping dynamic brain connectivity using a trial-based TE analysis of EEG source data, and the application of this technique to the investigation of semantic and number processing within the brain. The first paper (Chapter 2) documents the analyses of a semantic category and number magnitude judgement task using traditional ERP techniques. As predicted, the well-known semantic N400 component was found, and localised to left ATL and inferior frontal cortex. An N365 component related to number magnitude judgement was localised to right superior parietal regions including the IPS. These results offer support for the hub-and-spoke model of semantics, and the triple parietal model of number processing. The second paper (Chapter 3) documents an analysis of the same data with the new trial-based TE analysis. Word and number data were analysed at 0-200ms, 200-400ms, and 400-600ms following stimulus presentation. In the earliest window, information exchange was occurring predominately between occipital sources, but by the latest window it had become spread out across the brain. Task-dependent differences of regional information exchange revealed that temporal sources were sending more information to occipital sources following words at 0-200ms. Furthermore, the direction and timing of information movement within a front-temporal-parietal network was identified during 0-400ms of the number magnitude judgment. The final paper (Chapter 4), documents an attempt to track the influence of TMS through the brain using the TE analysis. TMS was applied to bilateral ATL and IPS because they are both important hubs in the brain networks that support semantic and number processing respectively. Left ATL TMS influenced sources located primarily in wide-spread left temporal lobe, and inferior frontal and inferior occipital cortices. The anatomical connectivity profile of the temporal lobe suggests that these are all plausible locations, and they exhibited excellent spatial similarities to the results of neuroimaging experiments that probed semantic knowledge. The analysis of right ATL TMS obtained a mirror image of the left. Left parietal stimulation resulted in a bilateral parietal, superior occipital, and superior prefrontal influence, which extended slightly further in the ipsilateral hemisphere to stimulation site. A result made possible by the short association and callosal fibres that connect these areas. Again, the results at the contralateral site were a virtual mirror image. The thesis concludes with a review of the experimental findings, and a discussion of methodological issues still to be resolved, ideas for extensions to the method, and the broader implications of the method on connectivity research.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright Statement

I. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

II. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

III. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

IV. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property University IP Policy (see http://documents.manchester.ac.uk/display.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses.
Contributions of Co-Authors

Chapters 2, 3, and 4 are presented in the form of papers, which were all co-authored by Dr. Marcelo Montemurro, Dr. Gorana Pobric, and Dr. Stephen Welbourne. Co-authors provided guidance on writing, formatting, and editing these papers and the other chapters as well. Data acquisition and analyses were performed by the author following direction from all co-authors.
Acknowledgments

I would like to thank my supervisors Dr. Marcelo Montemurro and Dr. Gorana Probric, plus my original principle supervisor Dr. Stephen Welbourne, for their help, support, and, most importantly, patience during the project. I’m sure I’ve not been the easiest student to manage, and I am grateful that you were always willing to meet and discuss any problems (real or imagined) that I had encountered. Without you all I would never have finished.

A big thank you to all my colleagues in the Neuroscience and Aphasia Research Unit who helped to make the day-to-day grind of a PhD bearable: Dr. Claude Bajada for discussions of all things MATLAB; Dr. Ajay Halai and Dr. Jen Gethin for EEG methods help; and the residents of Room T4 (past and present) for providing a fun working environment. I would also like to thank my advisors, Dr. Karl Embleton and Dr. Paul Warren, and the admin staff of the Faculty of Biology, Medicine and Health for answering queries that I probably should have known the answers to already.

I must of course extend my thanks to the Biotechnology and Biological Sciences Research Council who provided the funding for my PhD, and to University of Manchester staff who manage the Doctoral Training Partnerships.

Last, and by no means least, I would like to thank all my family and friends for their love and support. To Kathleen: thanks for providing the calmness needed to keep me (mostly) sane throughout the last 4 years. To Lynda: thanks for your unstinting encouragement. To Lucas: your arrival made the journey harder, but it has made the ending even sweeter.

For Edwin Stuart and Lucas Edwin
Thesis Overview

This thesis is written in the alternative format. Chapters 2, 3 and 4 are written as paper submissions with their own introductions and discussions. Consequently, there will be some unavoidable similarity in these chapters. Chapter 1 is an introduction providing a broad overview of the theoretical background and rationale for the thesis. Chapter 5 reviews the findings from the papers and highlights some outstanding methodological issues and possible future work.
Chapter 1 – General Introduction

The earliest investigations of brain function were neuropsychological studies that examined how lesions, or atrophy, at specific regions affected cognitive behaviours in both humans and animals. The conceptual framework employed by neuropsychology was localisational in nature, and posited that a particular area of the brain had a unique function (Friston, 2011). This localisational perspective of brain function can be traced back to the development of Phrenology by Gall in the late 18th century. However, by the end of the 19th century some scientists were beginning to argue for a distributional perspective of brain function which proposed that cognitive functions were underpinned by networks of several connected brain regions (Phillips et al., 1984). This distributional perspective can be traced back to the work of Wernicke (1874, cited in Catani & Ffytche, 2005) who helped to popularise the concept of disconnection syndromes, which are characterised by a breakdown of communication between healthy cortical regions produced by lesions in the white-matter tracts (Catani and Mesulam, 2008). For the first half of the 20th century, the localisational perspective held sway due to complementary results obtained from excitation and ablation studies, and some evidence from pathology (Phillips et al., 1984). However, the influential publication ‘Disconnexion syndromes in animals and man’ (Geschwind, 1965) helped to re-establish the distributional perspective by arguing that brain function was the product of wide-spread brain networks. Towards the end of the 20th century, an ever-increasing number of neuroimaging and analysis techniques were utilised to investigate healthy and disordered brain function for the first time in vivo. This neuroimaging evidence revealed that cognitive functions seemed to be underpinned by activity in multiple brain regions, and offered strong support for the distributional perspective (Martin, 2007). Cognitive neuroscience has since adopted the twin concepts of functional specialisation and integration to explain how distributed brain regions combined to produce cognitive function (Stephan et al., 2015). Functional specialisation suggests that individual brain regions perform certain computations, and functional integration suggests that the output of these computations is informed by the
context-dependent inputs from connected brain areas. Cognitive function arises from interactions between brain regions as the inputs they receive, and the outputs they produce, flow around the network. In short, you cannot investigate brain function without considering connectivity.

Within cognitive neuroscience, an analytical framework of connectivity identifies three key distinctions between anatomical, functional, and effective connectivity (Friston, 1994). Anatomical connectivity (AC) is defined as a complete description of the connections that exist between populations of neurons. Ideally, AC should include the gross structural features and the biophysical components of these connections. However, for the sake of simplicity, most neuroimaging studies define AC as the physical fibre tracts that connect distant regions of the brain (Sakkalis, 2011). Functional connectivity (FC) describes a pattern of synchronised neuroimaging signal dynamics (Lee et al., 2003). FC is defined as a statistical dependency between the activities of distinct neuronal populations (Friston, 2011). Effective connectivity (EC) is defined as the influence that one neural system exerts over another (Friston, 2011). Using quantifiers of EC, a particular region of the brain could be found to drive the activity within distant regions during cognitive processing, and this influence could be direct or indirect depending on the anatomic connections between them (Friston, 1994). Essentially, FC and EC are two ways of characterising functional integration.

This thesis documents a body of research that had two main aims: (1) to develop and test a new method of inferring effective connectivity within EEG data; (2) to use the method to investigate semantic and number processing. This introductory chapter is divided into six sections. The first discusses some of the most popular imaging and stimulation methods used to map brain function including their suitability for investigating connectivity. I will also review the popular methods of combining them to bypass their inherent limitations. Brain mapping techniques are used to obtain a signal of cortical dynamics that can be studied using a variety of different connectivity analyses. In section two, I will discuss the strengths and weaknesses of some popular methods used
to quantify connectivity within electrophysiological data. Some analysis methods are purely correlational and can’t be used to investigate effective connectivity, whilst others rely on the prior specification of connectivity models so are unsuitable for exploratory analyses. In section three, I will introduce Transfer Entropy (TE) which is the analysis technique used to quantify connectivity by the new method. TE is a non-parametric statistic that can measure the direction of information transfer between signals, and is suitable for data-driven investigations of effective connectivity. Quantifiers of connectivity have historically been applied to sensor-level electrophysiological data, but there is a convincing argument that states they should be applied at the source-level instead. There are many techniques that can be used to produce source reconstructions of electrophysiological data, and in the fourth section I will discuss them. The secondary aim of the thesis was to investigate whether the new TE analysis could be used to map connectivity between the brain regions that support semantic and number processing. Section five summarises the evidence which suggests that both these important cognitive functions are underpinned by widely distributed brain networks. Finally, I will sum up the open issues and aims of the thesis before briefly summarising the rest of the thesis chapters.

1.1 Brain-mapping techniques

Cognitive neuroscience employs several different imaging and stimulation techniques to map brain function and connectivity. Neuroimaging techniques can generally be separated into two types: indirect techniques, such as fMRI and PET, which are sensitive to various haemodynamic or metabolic processes associated with neural activation; and electrophysiological techniques, including MEG and EEG, that directly record brain activity. Neuro-stimulation techniques, such as TMS, are used to perturb the brain, and investigate how the evoked activity affects the site of stimulation and any connected
regions. In the following section I will examine the strengths and weaknesses of these methods in relation to connectivity research.

1.1.1 Haemodynamic and metabolic methods
During brain function, metabolism rates change due to neuronal firing, and blood flow to these areas also changes due to respiratory demands. This increase, or decrease, in neural activity can be detected indirectly by neuroimaging methods such as single photon emission computerised tomography (SPECT), and positron emission tomography (PET). Currently, the most popular indirect neuroimaging method is functional magnetic resonance imaging (fMRI) because it is quicker, cheaper, and does not require the intravenous administration of radioactive materials unlike the alternatives. fMRI determines neuronal activity indirectly by quantifying the amount of deoxyhaemoglobin in the brain via the blood oxygen level dependent (BOLD) haemodynamic response. The BOLD response has been found to correlate with local field potentials, so BOLD reflects the synaptic input activity of a neuronal population, rather than its firing output (Logothetis et al., 2001). When neurons become active, the vascular system increases blood flow to the required area, and an abundance of oxygenated haemoglobin builds up in the locality. This leads to a relative reduction in the concentration of deoxyhaemoglobin at this area, and a corresponding increase of signal in the resulting MR images (Amaro and Barker, 2006). fMRI spatial resolution depends on the magnetic strength of the fields employed, but it can typically be down to the range of millimetres. However, the temporal resolution of fMRI is relatively poor because the vascular response behaves like a filter which blurs the temporal information of the underlying neural activity. Furthermore, inferring connectivity from fMRI data is problematic because the relationship between neural activity and the BOLD signal is not linear (Logothetis et al., 2001). In summary, both PET and fMRI provide whole-brain coverage with excellent spatial resolution, but their temporal resolutions are too slow to capture the transient patterns and frequency specific nature of neuronal interactions (Gross et al., 2001).
Consequently, indirect methods such as fMRI are unsuitable for the investigation of dynamic connectivity.

1.1.2 Electrophysiological methods

Electroencephalography (EEG) measures the internal electrical activity of the brain via a network of electrodes positioned on the scalp (Baillet et al., 2001), whereas magnetoencephalography (MEG) records the magnetic field fluctuations induced by electrical brain activity using gradiometers or magnetometers positioned near the scalp (Cohen, 1968). The neuronal activity recorded by EEG and MEG (M/EEG) is mostly generated by the excitatory and inhibitory post synaptic potentials of pyramidal neurons located in layers III, V, and VI of the cortex. An important difference between EEG and MEG is the fact that EEG recordings are dominated by the activity in neurons arranged perpendicular to the plane of the scalp, whereas MEG signals are produced by neurons arranged tangentially. Consequently, EEG is more sensitive to neurons located in gyri, and MEG to those found in sulci (Baillet, Mosher, & Leahy, 2001; Ferree & Nunez, 2007). The temporal resolution of M/EEG is only limited by sampling frequency (≤ 1ms).

However, electrophysiological methods do possess spatial resolutions that are relatively poor when compared to the haemodynamic/metabolic techniques. Calculations based upon the current density of cortex, and the empirical measurements of current generation by neurons suggest that the magnitude of electrical activity recorded by M/EEG could be generated by areas of cortex that are 5mm² (Baillet et al., 2001). Despite this methodological weakness, the excellent temporal resolutions of electrophysiological methods mean they are able to capture the patterns of neuronal coupling (Gross et al., 2001), and as such they are optimal for investigating cortical connectivity.
1.1.3 Transcranial Magnetic Stimulation

Transcranial Magnetic Stimulation (TMS) can be used to non-invasively evoke neural activity in cortical regions via electromagnetic induction (Barker et al., 1985). It is unique amongst brain mapping techniques because it can be used to cause transient and reversible effects, which can lead to an inhibition or facilitation of cognitive function, with high temporal resolution. This effect is often referred to as a “virtual lesion” (Walsh and Cowey, 2000). Typically, the induced electric field rises to a maximum and reverses back towards zero in about 200μs, which leads to highly synchronous neuronal activation (Siebner, et al., 2009). The size, shape, and depth of the induced electric field are modulated by the intensity and frequency of the TMS pulse, the anatomy of the target region, and the angle of the coil relative to the surface of the skull. Figure-of-eight shaped coils can be used to improve the spatial resolution of TMS. However, even though a small bit of cortex might be in the peak of the magnetic field, 100-200mm² of cortical surface can be affected (Wagner et al., 2009). Only surface areas of cortex can be stimulated with any confidence because the magnetic field intensities dissipate rapidly as they move away from the coil. Theoretically, regions located in deep sulci, or the medial temporal lobes, could be targeted using higher intensity pulses, however any cognitive effects observed would be difficult to assign to the deep-lying site alone as the superficial regions would be strongly activated as the field passed through (Siebner et al., 2009).

Investigations of connectivity have been performed using dual site TMS paradigms (dsTMS), whereby the effect of TMS at a target region is modulated by a conditioning pulse of TMS applied to a connected region. The influence of the conditioning pulse can be investigated at different time points during cognitive processing thanks to the temporal resolution of TMS (O’Shea et al., 2008). However, the usefulness of dsTMS connectivity investigations are confined to motor and visual networks because the influence is quantified by its effect on any subsequent motor-evoked or phosphene-evoked potential (Angstmann and Siebner, 2012, p78). Consequently, TMS must be combined with other brain mapping techniques to investigate connectivity at brain areas that do not produce physical manifestations of the stimulation influence.
1.1.4 Combining brain mapping techniques

The combining of neuroimaging techniques is a popular theme within connectivity research. By combining neuroimaging modalities, the inherent limitations of each technique can be overcome, and data across a wide variety of temporal and spatial resolutions can be used to inform experimental conclusions and hypotheses (O’Shea et al., 2008; Rykhlevskaia et al., 2008). EEG-fMRI has the potential to combine the spatial resolution of fMRI and the temporal resolution of EEG (Ritter and Villringer, 2006). However, EEG and fMRI are sensitive to different temporal and spatial scales, and they also quantify different types of neural activity: fMRI measures synaptic inputs (Logothetis et al., 2001); and EEG is sensitive to post synaptic potentials of pyramidal neurons (Baillet et al., 2001). Therefore, co-registration of the signals can prove problematic because their generators do not always overlap (Nunez and Silberstein, 2000). By combining TMS and fMRI it is possible to use the spatial resolution of the latter to map the effects of evoked neural activity at brain regions that are connected to the site of stimulation (Ruff et al., 2009). However, the TMS pulse is known to cause distortions and signal loss in MRI images (Baudewig et al., 2000), and the slow temporal resolution of fMRI makes it impossible to characterise the nature of dynamic interactions between connected brain regions (O’Shea et al., 2008). Due to the synergy in their fast temporal resolutions, combined TMS-EEG can provide a real-time picture of cortical reactivity and connectivity. TMS produces a consistent pattern of deflections, known as the TMS-evoked potential (TEP), in EEG data at areas of the brain that do not produce an overt physical marker of excitability (Ilmoniemi and Kicić, 2010; Miniussi and Thut, 2010).

Neuroimaging of connectivity has been a steadily growing research area over the last decade (Friston, 2011), and there are various connectivity analyses that can be applied to neuroimaging and neuro-stimulation data. As discussed, the excellent temporal resolutions of electrophysiological imaging modalities are optimal for investigating the transient, frequency specific, interactions that characterise brain function. Therefore, we wanted to produce a method that analysed MEG or EEG data. MEG possess a better
spatial resolution than EEG (Cohen and Cuffin, 1983), but systems are not accessible to most researchers because they are expensive to purchase and maintain. We wanted our new method to be practical, cheap, and easily executed with equipment that is readily available at most research facilities. Therefore, an EEG-based method was the sensible option. In the next section, I will discuss some popular analysis techniques used to quantify connectivity within EEG data.

1.2 Quantification of functional and effective connectivity

Quantifiers of FC can be differentiated into methods that search for linear or non-linear correlations, and further still into time- or frequency-dependent procedures. EC methods can be grouped into two broad classes: those based on process models; and those that are data-driven. The following review section is by no means exhaustive, and there are more comprehensive reviews available for the interested reader (Pereda et al., 2005; Greenblatt et al., 2012).

1.2.1 Correlation/Covariance

The simplest quantification of FC measures the cross-correlation or linear covariance between the amplitudes of signals of electrode pairs as a function of time (Brazier and Casby, 1952). Linear covariance in the time domain has been extended to produce metrics that were sensitive to non-linear dependencies within EEG data (Quian Quiroga et al., 2002). However, correlation analysis has two major drawbacks: (1) volume conduction effects and overlapping sensor fields can artificially boost correlations between electrodes; and (2) causality cannot be inferred via correlation, so the direction of information flow between signals cannot be determined (Greenblatt et al., 2012). Time domain correlation/covariance has largely been superseded by quantifiers of FC in the frequency domain because they represent the easiest method available to study the synchronous oscillations that are thought to underpin FC (Varela et al., 2001; Fries, 2005).
1.2.2 Coherence

Coherence quantifies the extent of synchrony between frequency components of EEG data recorded at two electrodes (Pivik et al., 1993). Essentially, coherence measures the covariance of frequency between two signals, and is calculated as the cross-spectral density function between the two signals normalised by their individual auto-spectral density functions. The magnitude of coherence between two signals at a frequency of interest ranges from 0 to 1: 0 indicates that activities between each signal in this frequency are linearly independent, and 1 indicates that they are 100% interdependent (Pereda et al., 2005). Like time-based correlation analyses, coherence is especially susceptible to volume conduction effects. This problem can be further compounded by the choice of reference electrode; if the reference electrode is located near an experimentally active brain region (relative to the paradigm), then the coherence function might interpret this common source as some interdependence between the pair of electrodes that is not present (Pereda et al., 2005). Solutions to the problem of volume conduction have been proposed including the use of the imaginary part of coherence (Nolte et al., 2004), the phase lag index (Stam et al., 2007), and the phase slope index (Nolte et al., 2008). Coherence estimates can also be affected by changes in amplitude, so signals at electrodes that are correlated in phase, but not amplitude, might be missed by classical coherence (Pereda et al., 2005; Greenblatt et al., 2012). To avoid this obstacle, non-linear measures of frequency dependence have been formulated.

1.2.3 Phase synchronisation

Phase synchronisation (PS) measures are used to find non-linear interdependences in the frequency domain within EEG data. PS refers to a state where two oscillating systems are synchronised in the frequency domain, but their amplitudes may remain uncorrelated (Pikovsky, Rosenblum, & Kurths, 2003, p20-21). PS measures are typically applied between pairs of EEG electrodes, and they look for correlations between the signal frequencies irrespective of their voltage amplitudes (Pereda et al., 2005). The first step in any PS analysis of electrode signals is to extract the phase, a popular index of PS
known as the phase locking value (PLV) (Lachaux et al., 1999) utilises a wavelet transform to accomplish this. Phase values at a number of time points within the signals of interest are extracted, and the relative phase difference is calculated between the phases at both signals for each of the time points. PLV can be thought of as the average phase difference across all these time points.

1.2.4 Mutual Information

Mutual information (MI) differs from the measures of FC discussed above, as it is based on concepts derived from information theory. To calculate MI between two signals, Shannon entropies must first be estimated from the data. The Shannon entropy is a fundamental metric within information theory which quantifies the uncertainty (or variability) in a signal (Shannon, 1948). The Shannon Entropy ($H$) of a discrete random variable $Y$, with a probability distribution $P(y)$, is defined by equation 1.1. If all values of $Y$ are equally probable, then there is maximum uncertainty in the system and the entropy is maximised (Pereda et al., 2005).

$$H(Y) = - \sum P(y) \log_2 P(y)$$  \hspace{1cm} (1.1)

MI (equation 1.2) extends the Shannon entropy to quantify the reduction in uncertainty of a particular signal when observations of another, potentially coupled, signal are incorporated into the measurement (Vicente et al., 2011). As in equation 1.1, $H(Y)$ is the Shannon entropy of signal $Y$, and $H(Y|X)$ is the Shannon entropy of signal $Y$ when it has been conditioned to signal $X$.

$$MI_{XY} = H(Y) - H(Y|X)$$  \hspace{1cm} (1.2)
When applied to EEG data, MI quantifies how much extra information can be determined for the data at electrode Y, when the data at electrode X are also taken into consideration. $MI_{XY}$ will equal 0 if the time-courses recorded at each electrode are independent; no reduction in uncertainty at electrode Y is gained from knowledge of data at electrode X. However, $MI_{XY}$ will be positive if there is interdependence and rise to a maximal value equal to $I_Y$ for identical time-courses (Pereda et al., 2005). MI is sensitive to linear and non-linear interactions (Schoffelen and Gross, 2009), but the direction of influence between the two time-courses cannot be inferred because MI is a symmetric measure ($MI_{XY} = MI_{YX}$) (Vicente et al., 2011). To try and establish a level of asymmetry (and by extension, build up a picture of causal influences), lagged MI was introduced to account for the theoretical delay in influence from one electrode to another (Greenblatt et al., 2012). However, this method can result in the declaration of spurious causal interactions between pairs of electrodes due to shared information from a third source (Schreiber, 2000). Furthermore, it is often difficult to obtain a reliable measurement of MI between EEG sensors as a non-biased estimation requires a lot of data. This constraint is at odds with an event-related potential (ERP) analysis of a cognitive experiment because the continuous data are split into small time intervals (epochs) that represent the data pre- and post-stimulus presentation (Pereda et al., 2005).

Following on from these correlational measures of FC, I will now discuss quantifiers of EC that are applied to EEG data. These methods can be used to determine causality within connectivity networks so that the direction of influence can be revealed.

1.2.5 Dynamic Causal Modelling

Dynamic Causal Modelling (DCM) is process model-based quantifier of EC that was first developed for fMRI data analysis, but it has been extended to investigate EEG data too (Kiebel et al., 2009). In a standard DCM analysis, multiple network models are specified, as a set of mathematical sources (neural mass equations) and their interactions (differential equations), and Bayesian inference is used to identify the model that best accounts for the observed data. The assumptions, parameters, and equations that
provide the foundation of DCM analysis have been derived from neurophysiological research. However, it is impossible to adequately represent neural activity as a set of mathematical equations; this is especially true of the fine-grained detail that EEG provides (Daunizeau et al., 2011). Furthermore, DCM requires prior knowledge about the input to any model tested and its influence on the connections, which can be problematic because a significant amount of brain activity is internally generated, and is known to affect how neuronal populations respond to stimuli (Arieli et al., 1996). Therefore, it can be argued that DCM is unsuitable for exploratory analyses (Vicente et al., 2011).

1.2.6 Granger causality measures

Granger causality (GC) refers to an operational definition of connectivity that can be measured using various methods. The past of a time series Y can provide information about its future. If this predictive information about Y’s future is improved by utilising the past of a second time series X as well, then X can be thought of as a Granger-cause of Y (Granger, 1969). Granger causality has been estimated between electrophysiological signal amplitudes using multivariate autoregressive models (MVAR) which allies it with process model methods of quantifying EC. Two popular methods for quantifying GC within the frequency domain are the directed transfer function (Kaminski and Blinowska, 1991) and partially directed coherence (Sameshima and Baccalá, 1999). Despite their popularity, MVAR quantifiers of connectivity are only capable of detecting linear relationships between electrophysiological signals (Greenblatt et al., 2012), which could be a problem when tracking information flow within EEG data because the processes that produce electrophysiological data, such as ion-channel regulation and action potential generation, are inherently non-linear (Vicente et al., 2011).

As we have seen, there are many quantifiers of connectivity that are applied to electrophysiological data, and all have their strengths and weaknesses. However, a new method of quantifying connectivity within EEG data should be based on an analysis technique that is: (1) suitable for exploratory investigations because it makes no assumptions about the data; (2) can detect linear and non-linear effective connectivity
across a range of interaction delays because activity between brain regions may involve multiple pathways; and (3) is robust against the effects of field spread (Vicente et al., 2011). One such method is transfer entropy which provided the foundation of the new analysis method, and will be introduced in the next section.

1.3 Information Theory and Transfer Entropy

Transfer entropy (TE) is a non-parametric statistic that measures the directed transfer of information between two signals (Schreiber, 2000). In order to estimate the strength and direction of connectivity between signals, one must first estimate each signal’s Shannon entropy (Shannon, 1948), which is the cornerstone of information theory.

1.3.1 Information Theory

Information theory is a mathematical framework that can be used to study the quantification, communication, and storage of information. The Shannon entropy, $H$, quantifies the information contained in a signal by measuring uncertainty. Information and uncertainty are terms used within information theory to describe a process that can produce symbols. For example, a coin can produce 2 symbols, but before we flip the coin we are uncertain whether it will land heads or tails. Once flipped, we see the outcome and have received some information. A decrease of uncertainty leads to an increase of information. Let $x$ be a value at any time point in a signal that occurs with probability $p(x)$. Uncertainty $h(x)$ now becomes:

$$h(x) = \log_2 \left( \frac{1}{p(x)} \right)$$

1.3 Information Theory and Transfer Entropy

Transfer entropy (TE) is a non-parametric statistic that measures the directed transfer of information between two signals (Schreiber, 2000). In order to estimate the strength and direction of connectivity between signals, one must first estimate each signal’s Shannon entropy (Shannon, 1948), which is the cornerstone of information theory.

1.3.1 Information Theory

Information theory is a mathematical framework that can be used to study the quantification, communication, and storage of information. The Shannon entropy, $H$, quantifies the information contained in a signal by measuring uncertainty. Information and uncertainty are terms used within information theory to describe a process that can produce symbols. For example, a coin can produce 2 symbols, but before we flip the coin we are uncertain whether it will land heads or tails. Once flipped, we see the outcome and have received some information. A decrease of uncertainty leads to an increase of information. Let $x$ be a value at any time point in a signal that occurs with probability $p(x)$. Uncertainty $h(x)$ now becomes:

$$h(x) = \log_2 \left( \frac{1}{p(x)} \right)$$

1.3 Information Theory and Transfer Entropy

Transfer entropy (TE) is a non-parametric statistic that measures the directed transfer of information between two signals (Schreiber, 2000). In order to estimate the strength and direction of connectivity between signals, one must first estimate each signal’s Shannon entropy (Shannon, 1948), which is the cornerstone of information theory.

1.3.1 Information Theory

Information theory is a mathematical framework that can be used to study the quantification, communication, and storage of information. The Shannon entropy, $H$, quantifies the information contained in a signal by measuring uncertainty. Information and uncertainty are terms used within information theory to describe a process that can produce symbols. For example, a coin can produce 2 symbols, but before we flip the coin we are uncertain whether it will land heads or tails. Once flipped, we see the outcome and have received some information. A decrease of uncertainty leads to an increase of information. Let $x$ be a value at any time point in a signal that occurs with probability $p(x)$. Uncertainty $h(x)$ now becomes:

$$h(x) = \log_2 \left( \frac{1}{p(x)} \right)$$

1.3 Information Theory and Transfer Entropy

Transfer entropy (TE) is a non-parametric statistic that measures the directed transfer of information between two signals (Schreiber, 2000). In order to estimate the strength and direction of connectivity between signals, one must first estimate each signal’s Shannon entropy (Shannon, 1948), which is the cornerstone of information theory.

1.3.1 Information Theory

Information theory is a mathematical framework that can be used to study the quantification, communication, and storage of information. The Shannon entropy, $H$, quantifies the information contained in a signal by measuring uncertainty. Information and uncertainty are terms used within information theory to describe a process that can produce symbols. For example, a coin can produce 2 symbols, but before we flip the coin we are uncertain whether it will land heads or tails. Once flipped, we see the outcome and have received some information. A decrease of uncertainty leads to an increase of information. Let $x$ be a value at any time point in a signal that occurs with probability $p(x)$. Uncertainty $h(x)$ now becomes:

$$h(x) = \log_2 \left( \frac{1}{p(x)} \right)$$

1.3 Information Theory and Transfer Entropy

Transfer entropy (TE) is a non-parametric statistic that measures the directed transfer of information between two signals (Schreiber, 2000). In order to estimate the strength and direction of connectivity between signals, one must first estimate each signal’s Shannon entropy (Shannon, 1948), which is the cornerstone of information theory.

1.3.1 Information Theory

Information theory is a mathematical framework that can be used to study the quantification, communication, and storage of information. The Shannon entropy, $H$, quantifies the information contained in a signal by measuring uncertainty. Information and uncertainty are terms used within information theory to describe a process that can produce symbols. For example, a coin can produce 2 symbols, but before we flip the coin we are uncertain whether it will land heads or tails. Once flipped, we see the outcome and have received some information. A decrease of uncertainty leads to an increase of information. Let $x$ be a value at any time point in a signal that occurs with probability $p(x)$. Uncertainty $h(x)$ now becomes:

$$h(x) = \log_2 \left( \frac{1}{p(x)} \right)$$

1.3 Information Theory and Transfer Entropy

Transfer entropy (TE) is a non-parametric statistic that measures the directed transfer of information between two signals (Schreiber, 2000). In order to estimate the strength and direction of connectivity between signals, one must first estimate each signal’s Shannon entropy (Shannon, 1948), which is the cornerstone of information theory.

1.3.1 Information Theory

Information theory is a mathematical framework that can be used to study the quantification, communication, and storage of information. The Shannon entropy, $H$, quantifies the information contained in a signal by measuring uncertainty. Information and uncertainty are terms used within information theory to describe a process that can produce symbols. For example, a coin can produce 2 symbols, but before we flip the coin we are uncertain whether it will land heads or tails. Once flipped, we see the outcome and have received some information. A decrease of uncertainty leads to an increase of information. Let $x$ be a value at any time point in a signal that occurs with probability $p(x)$. Uncertainty $h(x)$ now becomes:

$$h(x) = \log_2 \left( \frac{1}{p(x)} \right)$$
Shannon championed the use of bits to measure information. In the case of a single coin toss, there are 2 possible outcomes, and we can record them as a 1 for heads and a 0 for tails. The uncertainty of a single coin toss is therefore $\log_2(2) = 1$ bit. If the coin was weighted to always land on heads, then we have no uncertainty about the outcome: $\log_2(1) = 0$ bits. The information in a system $X$ can be determined by calculating the entropy of the probability distribution of all its values. The Shannon Entropy is defined as the average uncertainty of the values $x$ weighted by the probability of their occurrence $p(x)$. If $X$ contains $M$ values, uncertainty $H(X)$ now becomes:

$$H(X) = - \sum_{x \in M} P(x) \log_2 P(x) \quad (1.4)$$

When the probability of each outcome is uniformly distributed, then uncertainty is maximised; there is less uncertainty in the system when a specific outcome is more likely to occur. Equation 1.4 represents the individual system entropy, but conditional entropies can be estimated from the data too. Equation 1.5 represents the Shannon entropy of system $X$ conditioned to system $Y$.

$$H(X|Y) = - \sum_{x} \sum_{y} P(x)P(x|y) \log_2 \left( \frac{P(x|y)}{P(x)} \right) \quad (1.5)$$
1.3.2 Transfer Entropy

Transfer entropy from signal $X$ to signal $Y$ is the increase of information gained about future values of $Y$ using the past of $X$ and $Y$. TE is calculated using equation 1.6:

$$TE_{X\rightarrow Y} = H(Y_t | Y_t^{(k)}) - H(Y_t | Y_t^{(k)}, X_t^{(k)})$$  \hspace{1cm} (1.6)

$H(Y_t | Y_t^{(k)})$ represents the Shannon entropy (reduction in uncertainty) of $Y$ at time $t$ when it has been conditioned (or fixed) to its own past $k$ time points. $H(Y_t | Y_t^{(k)}, X_t^{(k)})$ represents the Shannon entropy of $Y$ at time $t$ when it has been conditioned to $k$ past values of $X$ as well as $k$ past values of $Y$. TE is essentially an asymmetric extension of mutual information (Quian Quiroga et al., 2002). The TE framework utilises a Granger-causality definition of connectivity (Granger, 1969): a signal $X$ is said to cause signal $Y$, if the incorporation of past values of $X$ reduce the uncertainty in $Y$ more than the past values of $Y$ alone. The data at signals $X$ and $Y$ are first assumed to be Markov processes of order $k$. Therefore, the probability distribution of each signal at time $t$ is only dependent on $k$ past values, and the rest of the data is ignored (Figure 1.1). Intuitively, equation 1.6 can be thought of as the variability in $Y$ due to the past of $X$, the past of $Y$, and the rest of the system ($H(Y_t | Y_t^{(k)}, X_t^{(k)})$) subtracted from the variability in $Y$ due to the past of $Y$ and the rest of the system $H(Y_t | Y_t^{(k)})$. Therefore, $TE_{X\rightarrow Y}$ is equal to the variability in $Y$ due to the past of $X$, and if the value of $TE_{X\rightarrow Y}$ is positive then $X$ causes $Y$ to some extent. The conditioning of the Shannon entropies on the past values of each signal means that the metric is asymmetric ($TE_{X\rightarrow Y} \neq TE_{Y\rightarrow X}$). Therefore, a preferred direction of information can be determined between the signals (Pereda et al., 2005; Vicente et al., 2011). TE is suitable for exploratory research because it makes no assumptions of the data (Bressler and Seth, 2011). Furthermore, the estimation of TE positively correlates with gamma-band power in stimulated data (Buehlmann and Deco, 2010): a result that links TE to the theories of neural synchrony.
Figure 1.1. The principle of TE causality analysis: Sending signal $X$ causes receiving signal $Y$ if knowledge of $k$ past values of $X$ reduces the uncertainty in $Y$ at time-point $t$ more than using $k$ past values of $Y$ alone. 3 past time-points are used in the example.

### 1.3.3 Investigating effective connectivity within EEG data using Transfer Entropy

TE was used for quantifying EC within electrophysiological data soon after it was derived, and one of the first investigations was carried out on EEG data collected from epileptic patients (Palus et al., 2001). Long sections of continuous EEG were analysed with TE using a moving window, and periods of increased synchronisation between channel data and asymmetries of information flow were successfully located using the method. Both these types of transient phenomena are thought to pre-empt epileptic seizures, and it was concluded that TE is potentially useful for anticipating seizures. Despite this early success, TE has rarely been applied to EEG data since it was first derived. This was mostly due to the assumption that information theoretic analysis techniques require a relatively large number of data points so could not produce robust estimations from time-locked EEG data (Greenblatt et al., 2012). The suitability of TE to the analysis of EEG data was given a boost by the finding that it was able to detect different types of non-linear interactions within simulated EEG data (Vicente et al., 2011). Moreover findings
from a study indicate that the technique might be suitable for the analysis of averaged data in a fashion similar to a traditional ERP analysis (Martini et al., 2011). In this paper, participants were required to press a key with their right index finger when they saw a red circle on a screen, and their left when they saw a blue circle. A preferred direction of information flow was found from occipital electrodes to the rest of the brain at 100-200ms following stimulus presentation, and a similar flow of information from frontally located electrodes to the rest of the brain was found at 400-500ms. Therefore, TE detected dynamic patterns of information flow related to task performance, so could theoretically be used to map the brain networks associated with cognitive function.

Despite these encouraging results, TE has only been used to quantify EC within sensor-level EEG data, which compromises the conclusions that can be made about the cortical generators of any directed interactions identified. Indeed, this a limiting factor on the interpretation of all investigations of connectivity using EEG sensor-level data because it is susceptible to the problem of field spread (Nunez and Srinivasan, 2006, p85-86). Field spread can result in localised neural activity being represented at many sensors across the scalp. Indeed, applying connectivity analyses to data collected at the scalp tells us little about the interactions between underlying sources, and field spread has been shown to produce spurious correlations at spatial distant sensors in simulated EEG data, which can lead to an erroneous interpretation of connectivity (Schoffelen and Gross, 2009).

Furthermore, the interpretation of TMS-EEG experiments can be complicated by field spread because the location of TEP effects in scalp topography maps does not simply identify regions that are connected to the stimulation site: cortical sources that produced the TEP might not be located below the sensors. TE is able to mitigate the effects of field spread better than many quantifiers of connectivity because it utilises a time delay during estimation.

As previously stated, EEG possesses an optimal temporal resolution for the analysis of connectivity, but the problem of field spread in sensor-level data can affect both the spatial resolution and interpretation of connectivity analyses. To overcome these issues,
some researchers have argued that connectivity analyses should be applied to source reconstructions of electrophysiological data instead of the raw data recorded at the scalp (Schoffelen and Gross, 2009). In the next section I will summarise the most popular methods of performing EEG source reconstruction, which can be used to alleviate both these problems.

### 1.4 EEG source reconstruction

A lot of research has attempted to improve the poor spatial resolution of electrophysiological methods, and most it has focussed on two key areas: computing forward models and solving the inverse problem (see Grech et al., 2008; Hallez et al., 2007; Michel et al., 2004 for comprehensive reviews). The forward model incorporates tissue volume conductance and head geometry to describe the recorded scalp activity as a function of the internal sources (Mosher et al., 1999). The inverse solution attempts the opposite: modelling the source activity as a function of the electrode data. The inverse solution is ill-posed; this means that the electrical activity recorded at the scalp could be produced by multiple source configurations. The problem is derived from the fact that there are far more source configurations than independent observations. Furthermore, the solutions are unstable, so a small change in the scalp data can result in a large change to reconstruction (Grech et al., 2008). Despite this, reasonable solutions to the inverse problem can be achieved using a variety of techniques. Some of these techniques will be described below, as well as some common methods used to produce accurate forward models which are essential to the effectiveness of any inverse solution.
1.4.1 Forward Models

The simplest forward model is an extension of Ohm’s law to a 3D volume where the potential field $V$ recorded at $r$, generated by a unitary current dipole with a moment $d$ located at point $r'$ in an infinite, homogenous, and isotropic volume with conductivity $\sigma$:

$$V(r, r', d) = \frac{d \cdot (r - r')}{4\pi\sigma|r - r'|^3} \quad (1.7)$$

In one of the earliest papers to apply this technique, the 3D volume was a single sphere and the scalp potentials were modelled as a function of a single dipole located within the sphere (Frank, 1952). Modelling the head as single sphere was soon found to be too simplistic as the conductance of the skull is far lower than the scalp and brain. To account for these discrepancies, the head was modelled as 3 concentric spheres with appropriate conductance values (for an example see Salu et al., 1990). The discovery that the geometry of the head and the thickness and shape of the skull have a significant impact on forward solutions led to the rise in popularity of real head models (Hallez et al., 2007). The boundary element model (BEM) is a popular forward solution typically built from 3 tessellated surfaces that encapsulate the tissues of the head (Fuchs et al., 2002). These surfaces are produced using a segmented structural MR image. The 3D volume between each interface is assumed to be homogenous and exhibit isotropic conductance. Electrical potentials at the first boundary, typically housing the brain volume, are calculated in the centre of each triangle on the surface of the brain-skull interface. The potentials are then recalculated following the same trajectory that they would be travelling within a real head until the final solution is calculated at the boundary that represents the scalp. The BEM approach is popular within electrophysiological research because it is quick and easy to calculate, however the method does not account for the differences in conductivity between the tissues that form the head. Unlike BEM,
finite element modelling (FEM) utilises 3D volumes that are defined as a tetrahedral
mesh of small elements, and potentials are calculated throughout the volume rather than
just the interfaces (Wolters et al., 2004). Using this method, each individual element in
the volume can be assigned a conductance value, and is therefore a more accurate
representation of the conductance within a real head. However, the flexibility of the FEM
approach can also be a disadvantage as the parameters that need to be considered for
estimation are numerous. This can have significant impact on computation speeds, and
the validity of the forward model.

1.4.2 Inverse Solutions
The inverse solution can be calculated after successful estimation of the forward model.
Unsurprisingly, many methods have been proposed to solve the inverse problem, but
they can generally be placed into one of two groups: overdetermined and
underdetermined models. These groups are distinguished by the ratio of observations (or
electrodes in EEG) and estimated sources (Michel et al., 2004). Overdetermined models
estimate an inverse solution using fewer sources than the number of observations, and
they can be distinguished further into methods that utilise classical or Bayesian fitting to
determine the optimum model. Conversely, underdetermined models have more sources
than observations, and utilise deterministic approaches to find the optimum solution.
Techniques that fall into this category include minimum norm, LORETA, and
beamformers.

Overdetermined models assume that the EEG data recorded at the scalp can be modelled
successfully using a small number of dipoles. The inverse problem is ill-posed with an
infinite number of solutions, so the models must be constrained to enable the estimation
of an optimum solution. The most crucial constraint is the number of sources, and it is
usually informed by physiological and functional knowledge (Michel et al., 2004). Other
parameters used to constrain the models can include the strength, location, and
orientation of the dipoles (Grech et al., 2008). Classical fitting algorithms are utilised to
find the best solution by minimising some cost function between the real EEG data and
scalp potentials estimated using a forward model and multiple inverse models (see Wood, 1982 for an early example). However, it is inappropriate to assume that this solution is the actual configuration of sources that produced the data: in reality it is only slightly more likely to be correct than many other solutions (Michel et al., 2004). Furthermore, inverse solutions obtained using these algorithms can produce spurious source locations because they get stuck in areas of local minima where an increase in the cost function is produced by moving in any direction (Michel et al., 2004; Grech et al., 2008). Optimum inverse solutions using overdetermined models can also be found using Bayesian search algorithms. Unlike the classical fitting approaches that produce a single optimum solution, the Bayesian approach can be used to define many solutions that fit the observed data. Further interrogation of these models can reveal features of the sources that are highly likely to be accurate because they are shared across many solutions (Schmidt et al., 1999). Prior structural and functional information from modalities such as fMRI can also be incorporated into the model to improve the posterior distribution of solutions. While incorporating fMRI data into fitting algorithms can appear to be useful, it is not without its problems. The slow temporal resolution of fMRI might mean that some important sources are not identified because the transient activation did not produce a BOLD response. Furthermore, fMRI is tomographic so using structural data raises the chance of positioning sources in areas that are blind to electrophysiological recording. A fundamental criticism of underdetermined models is the fact that the exact number of sources cannot be known prior to analysis. Underdetermined models don’t make this assumption, which has led to a rise in their popularity.

Underdetermined (or distributed) models assume that sources of scalp potentials can be located at any of the points within a 3D grid overlaid on the brain. Unlike overdetermined models, the possible sources far outnumber the observations. Each grid point can be thought of as a fixed source whose orientation and strength can be altered to produce a solution that explains the scalp data. In the absence of any prior information, a unique source solution can be determined for any observed data using minimum norm
estimation (MNE: Hämäläinen & Ilmoniemi, 1994). This approach assumes that the
solution with the lowest Euclidean norm for a known potential field is the correct source
configuration. However, there is no physiological validity to this assumption. Indeed, the
algorithm favours solutions with small numbers of localised sources near the surface of
the brain, which can lead to the erroneous projection of deeper sources onto the surface
(Michel et al., 2004). Weighted techniques such as PROMS (Greenblatt, 1993), LORETA
(Pascual-Marqui et al., 1994) and FOCUSS (Gorodnitsky et al., 1995) have built on MNE
to compensate for this bias against deep sources, but these weights are purely
mathematical and have no physiological justification (Michel et al., 2004). Originally
developed for radar and sonar signal processing, beamformers can also be used to
determine the sources of electrophysiological data. Beamformers are spatial filters that
can be used to determine signal changes at a particular location over time. Unlike
overdetermined models, beamformers make no prior assumptions about the number or
location of sources. Instead, they scan a 3D grid of source points to determine the
contribution of each source to the observed data. Beamformers achieve the best
estimation of a target source’s signal by suppressing all the other sources. This ideal
beamformer could not be applied to electrophysiological data because neighbouring
sources in the brain often exhibit similar activity; it would only be useful if the scalp data
were produced by one source. Therefore, the spatial filter is adapted by attenuating the
suppression of sources other than the target. The linearly constrained minimum variance
(LCMV) beamformer (Van Veen et al., 1997) is a popular EEG source reconstruction
technique that applies adaptive filters in the time domain, and dynamic imaging of
coherent sources (DICS) has extended the LCMV beamformer into the frequency domain
(Gross et al., 2001).

Source reconstruction techniques, including the LCMV and DICS beamformers, have
drastically improved the spatial resolution of EEG. The knowledge of cortical generators
they provide has been added to research that utilises other neuroimaging techniques,
investigations of patients with neurological conditions, and neural network modelling to
inform the creation of models that try to explain the function of brain networks that underpin complex cognitive behaviours. In the next section I will briefly discuss the brain networks that support semantic and number processing which we investigated using the new method. These networks were chosen because they are well-researched and widely distributed throughout the brain, which would theoretically provide a good platform to test the new method.

1.5 Semantic and number processing

1.5.1 Semantic cognition in the brain

Human verbal and non-verbal behaviour is driven by multiple cognitive mechanisms that are collectively referred to as semantic cognition. Contemporary models agree that semantic cognition originates in a widely distributed network of sensorimotor regions that code the verbal and nonverbal information about objects (Martin, 2007; Patterson et al., 2007). For example, knowledge about colours activated the posterior fusiform gyrus (adjacent to regions known to process colour perception), and action knowledge produced neural activity in posterior middle temporal gyrus (Martin et al., 1995). However, the mechanism and location of the regions that support the formation of conceptual representations from sensory information is still a matter of debate. Models of semantic memory generally fall into one of two groups: distributed or hub-and–spoke models. Distributed models of semantic processing contend that conceptual representations arise through the amalgamation of neural activity between these modality-specific regions alone (Martin, 2007). In this model, the task requirements serve to direct the flow of neural activation around the network, and semantic associations between pairs of features (shape and name, or shape and action) are encoded along separate network connections. Whereas, hub-and-spoke models argue for the existence of a brain region (the “hub”) where associations between all feature representations are combined to form coherent semantic concepts (Patterson et al.,
In recent years, converging evidence from a number of different research areas has provided compelling support for a hub-and-spoke model of semantic processing with a hub located at the anterior temporal lobes (ATLs). Semantic dementia is characterised by relatively focal atrophy to polar and ventrolateral ATLs (Galton et al., 2001), and patients with this condition have profound difficulty performing tasks that probe semantic knowledge across all sensory domains (Patterson et al., 2007). Interestingly, similar semantic deficits were produced in healthy volunteers when repetitive transcranial magnetic stimulation (rTMS) was used to create a virtual lesion in left ATL (Pobric et al., 2007). Furthermore, neural activity at primary sensory areas was found to converge at left ATL around 400ms during a semantic category judgement task of written or spoken words (Marinkovic et al., 2003). This linked the ATL to the N400, which is a well-known ERP component that is commonly found during the processing of semantic anomalies (Kutas and Hillyard, 1980a). Moreover, computer models have shown that a hub region can be used to build up conceptual representations from different sensory inputs (McClelland and Rogers, 2003). Despite this wealth of evidence, semantic investigations using fMRI often showed a lack of activation in the ATLs which added weight to the alternative argument that the semantic hub does not exist (Martin, 2007). In recent years, it has been argued that this discrepancy between semantic dementia and healthy participants is a result of methodological problems with the fMRI technique: the ATLs are located near air-bone interfaces that are known to produce signal loss and image distortion. Indeed, a combined fMRI and PET study found that the same semantic category judgment task elicited anterior temporal lobe activity using both techniques, but it was far more wide-ranging and robust using PET which does not suffer from the same drop in signal-to-noise ratio (Devlin et al., 2000). Following the creation of semantic concepts, control processes in regions such as superior parietal and frontal lobes are then responsible for storing or manipulating these concepts in order to produce context-appropriate behaviours (Lambon Ralph, 2014).
In conclusion, semantic processing begins in primary sensory regions where feature extraction occurs and object attributes are encoded. At present, there is still debate about how this information is combined to form coherent concepts, but converging evidence highlights the importance of the ATLs in this role.

1.5.2 Number processing in the brain

The first investigations of brain function underpinning number processing were performed on patients suffering from acalculia. Acalculia is characterised by the inability to perform calculation tasks (Ardila and Rosselli, 2002), and is caused by lesions near the parieto-occipito-temporal junction or in the frontal lobe (Nieder and Dehaene, 2009). At the end of the 20th century, neuroimaging of mental arithmetic tasks revealed a consistent pattern of activity in bilateral parietal and prefrontal cortices (Dehaene et al., 1996; Rickard et al., 2000). The intra-parietal sulcus (IPS) seemed to be particularly important as it was reliably activated during tasks that probed mathematical operations (Chochon et al., 1999; Pinel et al., 2001). The IPS even responds to non-symbolic semantic representations of numbers such as a series of tones or a collection of dots (Castelli et al., 2006; Piazza et al., 2006). There is also evidence to suggest that parietal numerical processes can be found in pre-educated children (Cantlon et al., 2006; Izard et al., 2008), which supports the notion that the brain possess a language-independent quantification system (Nieder and Dehaene, 2009). In summary, the representation of symbolic and non-symbolic number processing seems to be supported by parietal and prefrontal regions, with a particular emphasis on the importance of the IPS.

A principle aim of this thesis is to document whether the new TE-based analysis can map task-dependent dynamic brain function. The networks that support semantic and number processing provide an excellent platform for achieving this aim because they are widely-distributed throughout the brain, perform multi-stage computations over long time-frames, and each possess a hub that is critical to network function: the ATL for semantics; and the IPS for numbers.
1.6 Aims of the thesis

This introduction gave a broad overview of the neuroimaging techniques used to investigate brain connectivity; analyses used to quantify connectivity including TE; source reconstructions methods applied to electrophysiological data; and a brief review of the semantic and number processing literature. The primary aim of this thesis is to document the creation of a new method capable of mapping dynamic brain connectivity. Electrophysiological neuroimaging techniques, such as EEG and MEG, were identified as being optimal for connectivity investigations (Gross et al., 2001). However, we wanted the new method to be easily executed with equipment possessed by most research facilities, so EEG was chosen as our data collection method. We also decided that the new analysis will utilise TE to characterise the patterns of brain connectivity because it is a model-free quantifier of linear and non-linear directed information exchange that can be used over a range of time-delays (Vicente et al., 2011). Inspired by arguments that suggest using EEG source data is preferential to channel data (Schoffelen and Gross, 2009), and the success of a time-locked TE analysis of EEG channel data (Martini et al., 2011), we decided to create a trial-based TE analysis of EEG source data.

The secondary aim of this thesis is to map the wide-spread networks that underpin semantic and number processing in the brain. To achieve this aim, we designed and executed a semantic category and number magnitude judgement task, and analysed the data with both traditional ERP techniques (Chapter 2) and the new trial-based TE analysis (Chapter 3). Furthermore, we wanted to ascertain whether the method was capable of tracking a TMS pulse from important nodes in the semantic and number processing networks to the brain regions with which they are functionally connected (Chapter 4). Perturbing the brain, using techniques such as TMS, is a less ambiguous method of assessing cortical connectivity.
Chapter 2

**ERP analyses of semantic category and number magnitude judgement tasks**

Chris Repper-Day, University of Manchester, UK
Marcelo A. Montemurro, University of Manchester, UK
Stephen R. Welbourne, University of Manchester, UK
Gorana Pobric, University of Manchester, UK
2.1 Abstract

The process of forming coherent semantic concepts is underpinned by a wide-ranging brain network with a hub at the anterior temporal lobes. However, one aspect of conceptual knowledge that does not seem to be dependent on the ATLs is knowledge about numbers. We investigated these processes concurrently by contrasting a task that probed object category knowledge with a task that probed number magnitude knowledge. Trials of both tasks consisted of a sequence of three priming stimuli and a fourth target. The fast temporal resolution of EEG enabled the investigations of task processing as the sequence progressed. As predicted, the well-known semantic N400 component was found in a cluster of midline fronto-central electrodes after all four word stimuli. ERP amplitudes reduced as the sequence progressed, which is a result common to many priming experiments that use implicit memory tasks. Sources of the N400 were localised to left ATL and inferior frontal cortex, and showed excellent convergence with PET and fMRI results. An N365 component related to number magnitude judgement was found in left lateralised parieto-occipital electrodes after the 2nd, 3rd, and 4th number stimuli, and was localised to right superior parietal regions including the IPS.
2.2 Introduction

Human verbal and non-verbal behaviour is often driven by multiple cognitive mechanisms that are collectively referred to as semantic cognition. An important facet of semantic cognition is semantic memory (or conceptual knowledge), which can be thought of as a database containing the information needed to bring meaning to everyday experience (Lambon Ralph, 2014). Most models of semantic memory agree that these concepts are formed from sensory information coded at modality-specific cortical regions; knowledge about an object’s colour, size, and use have been localised to individual brain areas (Thompson-Schill, 2003; Patterson et al., 2007). However, the topography of the networks that combine this object knowledge to form semantic concepts is still unclear. Distributed-only models argue that these sensory regions constitute the entire brain network (Thompson-Schill, 2003); whereas the distributed-plus-hub view contends that the network also contains a hub region where all of the sensory information is combined (Patterson et al., 2007). Recently, converging evidence from many research areas including neuropsychology, neuroimaging and neuro-stimulation studies, and computational modelling has provided strong support for models that locate the semantic hub in the anterior temporal lobes (ATLs). Semantic dementia is characterised by relatively focal atrophy to polar and ventrolateral ATLs (Galton et al., 2001), and patients with this condition have profound difficulty performing tasks that probe semantic knowledge across all sensory domains (Lambon Ralph, 2014). Evidence from studies of healthy participants using fMRI and PET (Visser et al., 2010), and TMS (Pobric et al., 2007) have also shown the importance of the ATLs to the formation of semantic concepts. Furthermore, computer models have demonstrated how information from sensory regions can converge at a hub to produce semantic concepts, and explain semantic task performance in both healthy and impaired participants (Rogers et al., 2004). One aspect of conceptual knowledge that does not seem to be dependent on the ATLs is number processing. Knowledge about numbers including size or whether a number is odd or even consistently activate the parietal cortex (Dehaene et al., 2003),
and neural activation in the region is often found during tests of mathematical
calculations (Hubbard et al., 2005; Nieder and Dehaene, 2009). Indeed, number
processing in patients with semantic dementia is often well preserved (Cappelletti et al.,
2012), which suggests that the semantic system possesses a dedicated mechanism for
numerical abilities.

2.2.1 Aims of the current study
In this study, we investigated the systems that support knowledge about objects and
knowledge about numbers using EEG. Specifically, we contrasted a task that probed
object category knowledge with a task that probed number magnitude knowledge. The
object category task was adapted for EEG from a PET and fMRI study that found
semantic activation in the ATLS (Devlin et al., 2000). This study confirmed that the ATLS
might be missing from some models of semantic processing because fMRI is often blind
to activity in brain regions near air-filled sinus/brain interfaces due to magnetic
susceptibility artefacts. In the task, a sequence of three prime words that belong to the
same object category (e.g. birds, vegetables etc.) was presented one stimulus at a time,
and participants had to decide whether a fourth word also belongs in the same category.
We chose to adapt this task for two reasons: (1) we would obtain converging evidence
from PET, fMRI and EEG using the same semantic task; (2) the high temporal resolution
of EEG would enable us to investigate semantic processing as the priming sequence
progressed. Our number magnitude judgement task used a similarly paced sequence of
three priming numbers, and a fourth target number that was either larger or smaller
than the final prime.

Semantic category judgement and number magnitude judgement processes are
supported by 2 large, wide-ranging, but crucially different, brain networks. We
confidently expected to find an N400 component during the category judgement decision
following word targets that we could locate to the ATLS. The N400 is a robust negativity
in the ERP/ERF waveform that is maximal at approximately 400ms following the
presentation of a stimulus. N400 have been found in semantic priming tasks using word,
picture, auditory, and video stimuli (Kutas and Federmeier, 2011). Semantic processing converges in the ATL and inferior prefrontal regions around 400ms after stimulus presentation (Marinkovic et al., 2003). During magnitude judgements, we predicted we would see an N400-like component, from sources located in superior parietal lobes, similar to those seen in other studies that probed numerical processing (Niedeggen and Rosler, 1999; Galfano et al., 2004). We also expected to find these components after some, or all, of the priming stimuli in both word and number sequences. Category judgment N400 might become more negative after each prime in the sequence and maximal after targets (Kutas and Hillyard, 1980b), or there could be a reduction in neural activation during the sequence as the same semantic processing is performed repeatedly (Henson and Rugg, 2003). We would not expect to find N400-like number magnitude components after the first prime because a relative magnitude judgment cannot be made until the second prime is presented.

2.3 Methods

2.3.1 Participants

16 volunteers (7 females and 9 males; mean age = 25.9 [SD = 5.5] years) were recruited to take part in the study. All participants were right-handed native speakers of English, had no history of neurological disorders, and possessed normal or corrected-to-normal vision. All participants gave informed consent, and the experiment was covered by the research group's ethical approval.

2.3.2 Stimuli

6 semantic categories were used: birds; fish; vegetables; clothes; tools; and vehicles. Each category word list contained 12 words (6/7 letters in length), they were drawn from the Hyperspace Analogue Language (HAL) corpus (Lund and Burgess, 1996), and they did not differ significantly in terms of frequency (F[5 66] = 0.16 , p = 0.976). Trial sequences were constructed so that all the words in the sequence were the same length,
and words were matched so that they appeared the same number of times across the 4 stimuli presentation points. Numbers were constructed of 6 or 7 digits. The magnitude change was created by increasing or decreasing the first digit by 1 across the first 3 stimuli presentations and the remaining digits in each number were randomly created. However, the first digit in the target number was always the same as stimulus 3, and the change in magnitude was created in the 2nd or 3rd digits within the target number. The remaining digits in the target stimulus were again selected randomly.

2.3.3 Experimental Paradigm

Trial sequences were presented on a 21” LCD monitor with a vertical refresh rate of 60Hz using the E-Prime 1.2 software (Psychology Software Tools, Pittsburgh, PA). In the semantic task, 4 word stimuli were presented in a sequence 1 word at a time. The first 3 stimuli belonged to the same semantic category, and the participant had to decide whether the fourth target stimulus could also be grouped in the same semantic category. The number magnitude judgment task utilised the same trial sequencing. The magnitude of the first 3 number stimuli increased or decreased relative to the preceding stimulus, and the participant had to assess whether the target number carried on the same trend (see Figure 2.1). Button responses, using a keyboard, were required from participants after semantically incongruent word targets and incongruent magnitude number targets only. All sequences were displayed in the centre of the screen in white 18-point text on a black background. There was a short practise block of 16 trials before the experiment began proper. Each trial commenced with a white fixation cross presented in the centre of the screen for 1000 milliseconds. Stimuli 1, 2, and 3 were presented for 500 milliseconds each and there was a 250-millisecond blank screen between consecutive stimulus presentations. The target stimulus was shown on screen for 2000 milliseconds. A complete trial lasted for 5250 milliseconds, there were 80 trials in each 7-minute experimental block, and there were 5 task blocks containing 200 word and 200 number trials in total. In both conditions, 160 trial sequences were congruent and required no
response from the participants, and 40 incongruent trials were used as catch trials. Participants were allowed to rest between blocks.

Figure 2.1. Word and number trial sequences: Pictorial representations of the appearance, order, and timing of the 4 sequence types (from top to bottom): 1.) Congruent tools; 2.) Incongruent birds; 3.) Congruent ascending; 4.) Incongruent descending.

2.3.4 EEG Data Acquisition

Continuous brain activity was recorded using an ActiveTwo Biosemi EEG system (Biosemi Inc., Amsterdam, The Netherlands). Data were recorded at 64 electrodes that were fixed to the scalp using an elastic cap studded with plastic electrode holders. The electrodes were arranged using an extended 10-20 layout. Electrooculogram data generated from blinks and eye movements were recorded from 4 electrodes placed on the face above and below each eye. All electrode impedances were maintained at less than 20 kOhm. The Biosemi system was connected to a computer running ActiView software (Biosemi Inc., Amsterdam, The Netherlands), and the data were digitised with a sampling rate of 512 Hz.
2.3.5 Data Pre-processing

EEG Data were pre-processed and analysed in MATLAB R2012a (MathWorks, Natick, MA) using the FieldTrip toolbox (version: 20140514; Oostenveld, Fries, Maris, & Schoffelen, 2011). Data were re-referenced to an average of all 64 electrodes, filtered from 1-60Hz, and visually inspected to remove trials that contained large artefacts resulting from involuntary muscle movements. Next, independent component analysis (RunICA: Bell and Sejnowski, 1995) was performed on each dataset, and components generated by pulse, eye blinks, and lateral eye movements were identified, from the time-course of their activity and location on the scalp, and removed. Bad channel data ($M = 0.929$ channels per participant, $SD = 0.997$) were repaired from neighbouring channel data using spherical spline interpolation. A final visual inspection was performed and any trials that still contained residual muscle artefacts were removed.

2.3.6 Data Analysis

Conventional task accuracy was calculated for congruent word and number trials, and average reaction times were calculated for incongruent word and number trials. ERPs were created following targets within word and number sequences for all electrodes from same category trials that elicited no response from the participant. A baseline of average activity 200ms before the presentation of stimulus 1 was used. To identify significant time windows in target data, a cluster-based (at least 2 electrodes) dependent-samples t-test with Monte Carlo randomization (1000 draws) was performed (Maris and Oostenveld, 2007). Using this method, significant differences in space and time can be identified within clusters of electrodes whilst controlling for multiple comparisons. The sum of the t-statistics within the cluster was used as the test statistic; if the probability of observing larger effects from shuffled data was below 5% then the cluster was assumed to be significant. Scalp topography maps of t-values during the significant time-windows were created and visually inspected to find significant electrodes. Average ERPs at these electrodes were created using word and number data to investigate the characteristics of the waveform following targets and the preceding 3 priming stimuli.
Finally, a linearly constrained minimum variance (LCMV) beamformer (Van Veen et al., 1997) was used to reconstruct the brain sources that contributed to the EEG data during the significant time-windows revealed by the cluster analyses. A head model was created using a boundary element model (BEM: Fuchs, Kastner, Wagner, Hawes, & Ebersole, 2002) that was derived from a segmentation of the Colin 27 stereotaxic template (Holmes et al., 1998). Next, a 5mm resolution regular source grid in MNI template space was created and inverse-warped to the Colin brain volume; the resulting model contained 8233 sources located inside the brain. The EEG electrodes were aligned to the scalp shell within the BEM, by transforming a Biosemi electrode layout template using the locations of the nasion, and left and right preauricular points defined on the Colin MRI. Components of interest were identified in the clustered ERP waveform, and a 100ms window of interest, centred at the maximum peak amplitude in both the word and number data, was used for the source reconstruction. Source contrasts of Words > Numbers and Numbers > Words during these time-windows were produced.

**2.4 Results**

**2.4.1 Behavioural Results**

Data from 14 participants are presented in this section because 1 participant was excluded due to poor performance on the behavioural task, and another because of a technical problem during EEG recording. Participants were paying attention and able to perform both tasks as word (M = 95.2%, SD = 3.4%; t(13) = 50.42, p < 0.001) and number accuracy (M = 89.9%, SD = 3.5%; t(13) = 42.02, p < 0.001) were significantly above 50% chance. However, there was a significant difference between word and number accuracy: t(13) = 6.9, p < 0.001.
There was also a significant difference between reaction times following incongruent word (M = 934.9ms, SD = 294.8ms), and number trials (M = 1034.3ms, SD = 295.2ms); t(439) = -5.336, p < 0.001. These results suggest that the participants found the number task harder than the semantic task.

2.4.2 Cluster Analysis Results
3 electrode clusters were found to be significant following targets. Figure 2.2 shows both the t-value scalp topography and the average ERP from the significant electrodes for each cluster. The earliest cluster was in right lateralised centro-parietal electrodes between 148-262ms following targets (p < 0.05). ERP within the first cluster of centro-parietal electrodes revealed that there was a larger N200, which peaked at 180ms, following number targets relative to words (Figure 2.2a). The second cluster was found between 336-496ms following targets in a left-lateralised cluster of parieto-occipital electrodes (p < 0.05). The ERP at the electrodes in this cluster exhibited a negative deflection following numbers targets only which peaked at 365ms. Henceforth we shall refer to this component as the N365 (Figure 2.2b). The final cluster was found in fronto-central electrodes between 348-422ms following targets (p < 0.01). ERP from these fronto-central electrodes revealed that there was an N400 component following word targets only with a maximum negativity at 395ms post-stimulus (Figure 2.2c).
Figure 2.2. T-value scalp topographies of significant clusters and average ERP after targets:
Significant electrode clusters revealed by the permutation analysis are coloured purple. Significant time-windows are coloured purple between (a) 148–262ms; (b) 336-469ms; and (c) 348-422ms.
2.4.3 Mean amplitude analysis of N400 and N365

As hypothesised, we found an N400 component following word targets. We also found an N365 component following number targets, which presumably represents some facet of the number magnitude judgement processing. Next, we investigated the ERP following presentation of the priming stimuli using the same clustered electrodes identified following targets. Negative components around 400ms after all 3 word primes were found (Figure 2.3), and there were negative components slightly earlier than 365ms following 2nd and 3rd number stimuli too (Figure 2.4). No component was visible in the waveform following the 1st number prime. Mean amplitude of these components within word and number ERPs were compared using a repeated measures ANOVA with factors of sequence type (word vs. number) and stimulus type (stimulus 1, 2, 3, and target). To account for latency variability, mean amplitude was calculated over a 100ms window that centred on the maximum value of the identified component after each stimulus.

At N400, ANOVA revealed that the main effect of sequence type (F[1, 13] = 25.558, p < 0.001) and stimulus (F[3, 39] = 10.383, p < 0.001) were both significant, but their interaction was not. Paired-samples t-tests revealed that there mean amplitudes were significantly larger after word targets (M = -0.817µv, SE = 0.164) than number targets (M = -0.395µv, SE = 0.102): t[13] = -4.85, p < 0.001. Furthermore, the amplitude differences between words and numbers at stimulus 1 (t[13] = -3.099, p < 0.01), stimulus 2 (t[13] = -3.860, p < 0.005), and stimulus 3 (t[13] = -5.088 , p < 0.001) were also significant (see the graph in Figure 2.3). The differences between mean amplitudes at each word presentation were also examined using post hoc tests. Mean amplitude at stimulus 1 (M = -0.781 µv, SE = 0.171) was significantly more negative than targets (M = -0.332, SE = 0.107): t[13] = -3.296, p < 0.01. Stimulus 1 was also significantly more negative than stimulus 2 (t[13] = -2.542, p < 0.05), and stimulus 3 (t[13] = -2.309, p < 0.05). No differences between mean amplitudes at other stimulus presentations reached significance.
Figure 2.3. Average ERP and mean amplitude of the N400 at all stimuli: Average ERPs within the fronto-central cluster following the 3 priming stimuli are shown. As indicated by the green arrows, N400 peaked at 379ms following stimulus 1, 359ms following stimulus 2, 375ms following stimulus 3, and 395ms following targets. Mean amplitudes were calculated using a 100ms window centred at these peaks. Error bars are ± 2 SE.

N365 was not present after stimulus 1, so the ANOVA factor of stimulus type only had 3 levels. ANOVA revealed a significant main effect of sequence type (F[1, 13] = 14.417, p < 0.005), but the main effect of stimulus presentation was not significant and neither was their interaction. Post hoc tests indicated that the difference between mean amplitude after word targets (M = 0.092µv, SE = 0.170) and number targets (M = -1.035µv, SE = 0.305) was significant (t[13] = 4.547, p = 0.001). Furthermore, mean amplitudes were significantly more negative for numbers than words following stimulus 2 (t[13] = 3.162, p < 0.01), and stimulus 3 (t[13] = 3.582, p < 0.005).
We hypothesised that there would be a component in the number data that represented the magnitude judgment, and that it would be temporally located around 400ms after targets. N365 was only found after stimulus 2, 3, and targets, and this is commensurate with the fact that relative magnitude judgements are only possible at these points in the sequence because there is no preceding number before stimulus 1. The t-value topography map at N365 (Figure 2.2b) indicated that left lateralised parietal-occipital electrodes were more negative than the same positions in the right hemisphere following number target during the significant time-window. To investigate this laterality effect further, we created ERP from left- and right-lateralised electrode clusters after each number presentation (Figure 2.5).
Interestingly, there appears to be an N2pc component in both the left and right ERP peaking between 225-275ms after each number stimulus. The N2pc is a subcomponent of the N200 family, and is thought to represent the degree of attention that is needed for processing stimuli features within the visual cortex (Patel and Azzam, 2005). Following the 1st stimulus there is no magnitude difference between left and right electrodes, but after 2nd, 3rd, and target stimuli it appears to be larger in the right hemisphere. Mean amplitudes for N365 were calculated for left- and right-lateralised electrodes using a 100ms window centred on the peak of the component. Paired samples t-tests revealed that the difference between mean amplitude after number targets in the left hemisphere (M = -1.396µv, SE = 0.414) and right hemisphere (M = -0.672µv, SE = 0.258) was
significant ($t[13] = 2.877, p < 0.02$). Furthermore, the amplitude difference at stimulus 2 ($t[13] = 3.029, p = 0.01$) and stimulus 3 ($t[13] = 3.753, p < 0.005$) were also significant.

### 2.4.4 Source Analysis

Bilateral sources in the $W > N$ contrast at N400 were found in anterior fusiform gyri, temporal poles, parahippocampal gyri, orbitofrontal cortices. Small clusters of sources were in the right cerebellum. Sources in the left hemisphere were found in anterior inferior and middle temporal gyri, and anterior prefrontal cortex. Bilateral sources in the $N > W$ contrast at N365 were found in posterior fusiform gyri, posterior inferior temporal lobe, precuneus, and cerebellum. Sources in the left hemisphere were found in postcentral gyrus. Sources in the right hemisphere were found in precentral gyrus, and inferior, middle, and superior occipital cortex. Furthermore, sources were found in right superior and inferior parietal lobule around the intraparietal sulcus (Figure 2.6).

**Figure 2.6. Source reconstructions of N365 and N400 components:** Source contrast of Numbers $>$ Words between 315-415ms (red), and Words $>$ Numbers between 350-450ms (green) following targets. Only top 10% of source magnitudes are shown.
2.5 Discussion

The principle aim of this research was to directly compare semantic processing of words and numbers within the same EEG experiment. We investigated the distinct networks that underpin the processing of semantic knowledge and number comparisons at millisecond time-scales using the high temporal resolution of EEG. By employing the same category judgement task previously used with PET and fMRI, our results will provide converging semantic evidence from a third neuroimaging technique. The task incorporated a lengthy priming sequence and a response to a target, so word and number processing could be analysed at multiple points during the task.

The first significant difference we found in the ERP was an N200 that was larger for numbers than words over right lateralised fronto-central, central, and centro-parietal electrodes. N200 with a central scalp topography like ours have been categorised as N2b, which are elicited by conscious stimulus attention (Patel and Azzam, 2005), and have amplitudes that are correlated with the difficulty of visual feature discrimination (Senkowski and Herrmann, 2002). This result supports the evidence from our behavioural data which suggests that correctly discriminating between a congruent and incongruent number target is more difficult than the semantic category judgement in word trials. The right lateralisation seen in the topography was repeated at the N2pc components visible in the ERP within parieto-occipital electrodes (Figure 2.5). N2pc is thought to represent the degree of attention that is needed for processing stimuli features within the visual cortex, and is larger in contralateral electrodes if stimuli are presented lateral to fixation (Luck et al., 1997; Folstein and Van Petten, 2008). This effect could have been caused by a systematic bias in our number stimuli which meant that participants only had to focus on the first 3 digits to perform the magnitude judgement. As the stimuli were centrally aligned, participants were presumably maintaining their attention on the left-hand side of the screen, which could result in the larger amplitude of the N2pc in right hemisphere electrodes.
To ascertain whether this lateralisation effect was caused by the location the magnitude change we would need to run the experiment again, and position the information required for a correct magnitude judgement in both visual fields.

2.5.1 Semantic category judgment

As hypothesised, we found an N400 component after all four stimuli in the word trials. Scalp topography maps revealed that the N400 response was located at fronto-central electrodes which is slightly at odds with the classic centro-parietal topography of N400 (Kutas and Hillyard, 1982). However, N400 have previously been found at frontal and central electrodes during a semantic memory search task that required participants to hold category labels in working memory (Mecklinger et al., 1992). Task performance in our experiment would require a similar working memory element that was not required in the original N400 studies. The largest negativity at N400 was found after the first stimuli in the sequence, and N400 amplitudes reduced as the trial continued. This result is common to many priming experiments that use implicit memory tasks: processing repeated stimuli results in reduced neural activity (Henson and Rugg, 2003). However, this is the first time that these repetition effects have been found processing repetitive semantic categories using EEG to the author’s knowledge.

The source reconstruction found semantic category task activation in bilateral orbitofrontal areas (BA 11 & 47), and left medial temporal lobe (BA 35), left temporal pole (BA 38), and left anterior inferior and middle temporal lobes (BA 20 & 21). This result is very similar to those obtained in the original PET/fMRI study using this task (Devlin et al., 2000). Furthermore, activation in these areas, amongst others, have been found previously when localising the sources of N400 using MEG (Halgren et al., 2002; Marinkovic et al., 2003), event-related optical signal (Tse et al., 2007), and intracranial electrodes (Halgren, Baudena, Heit, Clarke, & Marinkovic, 1994). Unlike Devlin and colleagues, we did not find activation at Broca’s area (BA 45), which is presumably related to the disparity of temporal resolutions between the indirect neuroimaging techniques they employed, and our direct measure of neural activity. We performed
source reconstruction on a window between 350-450ms following targets to produce our task activation images, but activity in Broca’s area has been isolated at ~300ms, before moving to ATL locations, when sentences with congruous/incongruous endings are read (Halgren et al., 2002). Therefore, task-related activation at Broca’s area will most likely have reduced during our time-window of interest.

2.5.2 Number magnitude judgment

As predicted, we located an N365 component associated with the magnitude judgement, but interestingly it was preceded by an N2pc component that was also affected by task performance. Following N2pc, both left and right ERP waveforms increased in polarity until the start of the negative deflection that forms the N365 in left electrodes only. This cross-over of left and right waveforms strongly suggests that the N365 is representative of different cognitive processes to those producing the N2pc: namely the number magnitude judgement. This component was in left lateralised parieto-occipital electrodes peaking at 365ms following number targets only. Further investigation revealed the same component was visible after the 2nd and 3rd stimuli as well, but not present after the 1st stimulus. This pattern of processing was expected as you cannot discern whether the number sequence is increasing or decreasing after the 1st stimulus. Laterality differences within number processing have been observed before, however the individual contribution of left and right hemispheres is still unknown (Andres et al., 2005). Furthermore, the laterality difference found in our experiment is hard to unpick from the effects of the systematic bias present in our stimuli. There were no significant differences in the N365 negativities as the sequence progressed towards targets, which would indicate that the mental effort expended at each magnitude judgement was similar.

Unlike semantic category activation, neural activity associated with the number magnitude task was mainly located in posterior regions. There was a large area of activation extending over inferior, middle, and superior occipital cortices (mainly BA 18). This pattern of activity could be associated with the right-lateralised N200 attention effect seen in the ERP that we have attributed to the systematic bias present in the
number stimuli. The most likely candidate for the origin of the N365 are bilateral precuneus and right superior and inferior parietal lobule around the intraparietal sulcus (BA 7): activation associated with number comparison tasks is often larger in the right hemisphere relative to the left (Dehaene et al., 2003). These regions are thought to be the location of the “quantity system” within the triple-code model of number processing (Dehaene et al., 2003). Furthermore, this area is activated in tasks that probe mental arithmetic (Chochon et al., 1999; Pesenti et al., 2000), and number comparison (Dehaene, 1996; Pinel et al., 2001).

2.5.3 Conclusion
Semantic concept formation and number magnitude calculation are carried out by widely-distributed brain networks. However, the topology of these networks and the operations they employ to perform these important tasks is still a matter of debate. Using the high temporal resolution of EEG, we could build a richer picture of brain function during performance of a semantic category judgement task that had previously been utilised in a PET/fMRI study. ERP analysis revealed the well-known semantic N400 component was present after all word stimuli in the priming sequence, but repeatedly viewing objects from the same semantic category resulted in response suppression and a reduction in neural activity. The cortical generators of the N400 were localised to regions in the left ATL and inferior frontal regions, which corresponded well with areas already identified in the original PET/fMRI study. The ERP analysis of number magnitude judgment found an N365 component which was largest in the ipsilateral hemisphere to the visual field that was being attended to. Source reconstruction of the cortical generators responsible for this component reaffirmed the importance of superior parietal regions, including the right intraparietal sulcus, to number magnitude processing.
Chapter 3

Trial-based transfer entropy analysis of semantic category and number magnitude judgement tasks

Chris Repper-Day, University of Manchester, UK
Gorana Pobric, University of Manchester, UK
Stephen R. Welbourne, University of Manchester, UK
Marcelo A. Montemurro, University of Manchester, UK
3.1 Abstract

EEG possesses the optimal temporal resolution to capture functional connectivity, but it lacks the spatial resolution to identify the active cortical locations. To circumvent this problem electrophysiological connectivity should be investigated at the source level. Transfer Entropy is a data-driven quantifier of connectivity that can infer the direction of linear and non-linear information exchange between signals over a range of time-delays. A novel trial-based TE analysis of EEG source data was used to map the brain networks that support semantic and number processing. We contrasted a task that probed object category knowledge with a task that probed number magnitude knowledge. Semantic word and number data were analysed at 0-200ms, 200-400ms, and 400-600ms following target presentation. In the earliest window, information exchange was occurring predominately between occipital sources, but by the latest window it had spread throughout the brain. Task-dependent differences of regional information exchange revealed that temporal sources were sending more information to occipital sources following words at 0-200ms. Occipital sources were sending more information to parietal sources following words at 0-200ms and 400-600ms, which could be evidence of separate feature extraction processes. Furthermore, the direction and timing of information movement within a front-temporal-parietal network was identified during the first 400ms of the number magnitude judgment.
3.2 Introduction

The development of neuroimaging techniques such as PET and fMRI, in the late 20th century, enabled the non-invasive study of brain function in healthy individuals for the first time. This methodological breakthrough led to an explosion of “Brain Mapping” research (Savoy, 2001). The results of this proliferation strengthened the argument that brain function was distributional in nature and underpinned by networks of several connected brain regions (Phillips et al., 1984). There followed an increase of research focussing on neural interactions between two or more brain regions, which gave rise to influential models of brain function that were underpinned by coherent neuronal activity (Singer, 1999; Varela et al., 2001; Buzsáki and Draguhn, 2004; Fries, 2005). These models argue that synchronous oscillations between disparate areas of the brain provide the means for neuronal populations to exchange and process information. In short, brain function originates from coordinated activity within and between connected areas of the brain.

3.2.1 The neuroimaging of connectivity

Cognitive neuroscience has adopted a conceptual framework which separates connectivity into 3 distinct types: anatomical, functional, and effective (Lee et al., 2003). Anatomical connectivity is defined as a complete description of the physical connections between brain regions, but is simplified to refer to the fibre tracts that connect distant areas (Sakkalis, 2011). These physical structures support the synchronous flow of neuronal information around functional brain networks. The coherent neural dynamics of a brain networks are described by the twin concepts of functional and effective connectivity (Friston, 2011). Functional connectivity (FC) describes patterns of statistical dependency between nodes of brain networks, and effective connectivity (EC) describes the directions of influence between the nodes as cognitive behaviours are performed.
Investigations of connectivity have been implemented with several neuroimaging techniques. For example, fMRI has provided a wealth of information about the locations of nodes within functional brain networks because it possesses exceptional spatial resolution (1-10mm). However, fMRI is an indirect measure of neural activity, so its temporal resolution is too slow to adequately map the transient interactions between network regions (Gross et al., 2001). Electrophysiological neuroimaging techniques, such as electroencephalography (EEG), do not possess this limitation as they are direct measures of brain activity. EEG possess a temporal resolution that is able to capture the neural dynamics at time-scales of less than 1ms, and as such it is optimal for quantifying FC and EC (Sakkalis, 2011). However, the spatial resolution of EEG is relatively poor due to the problem of field spread (also known as volume conduction), which means that brain activity can be recorded at more than one electrode at a time. In recent years, some researchers have suggested that connectivity measures should be applied to EEG source data, rather than channel data, in order to overcome this problem (Schoffelen and Gross, 2009).

3.2.2 Quantification of connectivity in EEG data

Quantifiers of functional and effective connectivity that are applied to EEG data can be separated into metrics that are sensitive to statistical dependencies in the time and frequency domain. Furthermore, some quantifiers can only detect linear relationships between signals whilst others are sensitive to nonlinear dependencies as well. The following section is intended as a brief review of FC and connectivity measures that are applied to EEG data (for more extensive reviews see (Pereda et al., 2005; Sakkalis, 2011; Greenblatt et al., 2012). The earliest connectivity studies simply used linear covariance between signal amplitudes as a function of time (Brazier and Casby, 1952), which was later extended to nonlinear dependencies (Quian Quiroga et al., 2002). Temporal covariance has largely been superseded by coherence, which is covariance in the frequency domain, because it is a simple way of determining whether two signals are functionally coupled at a particular frequency (Pfurtscheller and Andrew, 1999).
Covariance measures are popular because they are simple to implement. However, correlations between signals can be artificially inflated due to the overlapping of sensor fields (also known as volume conductance) which leads to spurious ‘connections’ being detected (Greenblatt et al., 2012). This problem has been counteracted using at least two methods: (1) the phase slope index factors out volume conductance effects by using the imaginary part of coherence only (Nolte et al., 2008); and (2) quantifiers of connectivity have been applied to source reconstructions of EEG data rather than the channel data (Schoffelen and Gross, 2009). Quantifiers of connectivity that utilise covariance can determine signals that are active at the same time, and are therefore functionally connected. However, it is impossible to determine any causality, and by extension EC, within this connection using instantaneous correlations. Quantifiers of EC in EEG data can be separated into model-based methods such as dynamic causal modelling (DCM: Kiebel, Garrido, Moran, Chen, & Friston, 2009), and model-free methods such as those based on Granger causality (GC: Granger, 1969). In DCM, multiple network models are specified, as a set of mathematical sources (neural mass equations) and their interactions (differential equations), and Bayesian inference is used to identify the model that best accounts for the observed data. However, because the models are specified a priori, DCM is not data-driven and is therefore unsuitable for exploratory analyses (Vicente et al., 2011). Granger causality refers to a definition of causality: a signal X is said to Granger-cause signal Y if the past values of X and Y predict future values of Y better than using the past of Y alone. Granger causality can be measured using a multivariate autoregressive model (MVAR), and two methods that utilise this approach to measure granger causality in the frequency domain are the directed transfer function (Kaminski and Blinowska, 1991) and partially directed coherence (Sameshima and Baccalá, 1999). MVAR models the EEG data as the output of a linear time-invariant system, so they are incapable of detecting nonlinear relationships within data (Greenblatt et al., 2012).
Many of the physiological processes that produce neural activity are inherently nonlinear (e.g. the action potential; regulation of ion channels etc.), so quantifiers of EC that are sensitive to nonlinear interactions are desirable: one such metric is transfer entropy.

### 3.2.3 Transfer entropy

TE measures the directed transfer of information between two signals using the Shannon entropy $H$ (Shannon, 1948), which is the cornerstone of information theory. The Shannon entropy quantifies the information contained in a signal by measuring uncertainty. A decrease of uncertainty leads to an increase of information. The information in a system $X$ can be determined by calculating the entropy of the probability distribution of all its values. The Shannon Entropy is defined as the average uncertainty of the values $x$ weighted by the probability of their occurrence $p(x)$. If $X$ contains $M$ values, uncertainty $H(X)$ now becomes:

$$H(X) = - \sum_{x \in M} P(x) \log_2 P(x)$$  \hspace{1cm} (3.1)

When the probability of each outcome is uniformly distributed, then uncertainty is maximised; there is less uncertainty in the system when a specific outcome is more likely to occur. Equation 3.1 represents the individual system entropy, but conditional entropies can be estimated from the data too. Equation 3.2 represents the Shannon entropy of system $X$ conditioned to system $Y$.

$$H(X|Y) = - \sum_{x} \sum_{y} P(x)P(x|y) \log_2 \left( \frac{P(x|y)}{P(x)} \right)$$ \hspace{1cm} (3.2)
Transfer entropy from signal $X$ to signal $Y$ is the increase of information gained about future values of $Y$ using the past of $X$ and $Y$. TE is calculated using equation 3.3. $H(Y_t \mid Y_t^{(k)})$ represents the Shannon entropy in signal $Y$ at time $t$ when it has been conditioned (or fixed) to its own past $k$ time points. $H(Y_t \mid Y_t^{(k)}, X_t^{(k)})$ represents the Shannon entropy of $Y$ at time $t$ when it has been conditioned to $k$ past values of $X$ as well as $k$ past values of $Y$ (Figure 3.1).

$$TE_{X \rightarrow Y} = H(Y_t \mid Y_t^{(k)}) - H(Y_t \mid Y_t^{(k)}, X_t^{(k)})$$ (3.3)

If the incorporation of past values of $X$ increases the information within $Y$ more than using the past of $Y$ alone, then $X$ is said to cause, or influence, $Y$. Intuitively, this equation can be thought of as the variability in $Y$ due to the past of $X$, the past of $Y$, and the rest of the system ($H(Y_t \mid Y_t^{(k)}, X_t^{(k)})$) subtracted from the variability in $Y$ due to the past of $Y$ and the rest of the system $H(Y_t \mid Y_t^{(k)})$. Therefore, $TE_{X \rightarrow Y}$ is equal to the variability in $Y$ due to the past of $X$, and if the value of $TE_{X \rightarrow Y}$ is positive then $X$ causes $Y$ to some extent. The conditioning of the Shannon entropies on the past values of each signal means that the metric is asymmetric ($TE_{X \rightarrow Y} \neq TE_{Y \rightarrow X}$), so a preferred direction of information can be determined between the signals (Pereda et al., 2005; Vicente et al., 2011). TE is suitable for exploratory investigations because it is model free. It has been used to detect different types of non-linear interactions within continuous data collected from epileptic patients (Palus et al., 2001), and continuous simulated EEG data (Vicente et al., 2011). Furthermore, the estimation of TE positively correlates with gamma-band power in simulated data (Buehlmann and Deco, 2010), which links TE to the theories of neural synchrony. Despite these positive results, TE has rarely been applied to EEG data since it was derived in 2000. This was mostly due to the assumption that information theoretic analysis techniques require a relatively large number of data points so could not produce robust estimations from time-locked EEG data (Greenblatt et al., 2012).
However, Martini et al. (2011) indicated that TE might be suitable for the analysis of averaged data in a fashion similar to traditional ERP analyses.

**Figure 3.1. The principle of TE causality analysis:** Sending signal X causes receiving signal Y if knowledge of k past values of X reduces the uncertainty in Y at time-point t more than using k past values of Y alone. 3 past time-points are used in the example.

### 3.2.4 Aims of the current study

The principle aim of the current study is to ascertain whether TE can be used define the patterns of dynamic information flow within time-locked EEG source data. In applying TE to source data, we propose to assess EEG connectivity whilst negating the problem of volume conductance. Dynamic imaging of coherent sources (DICS: Gross et al., 2001) has produced good results from source data using a measure of FC, but we hope that TE will produce a richer picture of network dynamics because it is sensitive to the direction of influence (EC). The data we have used were obtained from an EEG experiment in which the participants had to make semantic category judgements or number magnitude
judgements. The well-known semantic N400 component (Kutas and Hillyard, 1980a) was found in word data, and was localised to left inferior frontal and left anterior temporal lobes. The number magnitude judgement was associated with an N365 component that was localised to right superior parietal sources. Theoretically, TE could be used to elucidate the timing of interactions between the sources within these networks as a typical trial unfolds. We hypothesise that information from occipital sources should converge at the ATLs in semantic task data (Marinkovic et al., 2003), and in parietal lobes following number presentations at around 400ms (Nieder and Dehaene, 2009).

3.3 Methods

3.3.1 Participants

16 volunteers (7 females and 9 males; mean age = 25.9 [SD = 5.5] years) gave informed consent to take part in this study. They were all right-handed native speakers of English, possessed normal or corrected-to-normal vision, and had no history of neurological disorders. The study was approved by the local ethical committee.

3.3.2 Stimuli

A semantic category judgement task that was previously used in an fMRI experiment (Devlin et al., 2000) was modified for EEG. Semantic word stimuli from 6 categories (clothes, tools, vehicles, birds, fish, and vegetables) were drawn from the Hyperspace Analogue Language (HAL) corpus (Lund and Burgess, 1996). Each category contained 12 words (6x12 = 76 words total) with a length of 6 or 7 letters; word stimuli did not differ significantly in terms of frequency (F[5 66] = 0.16 , p = 0.976). In each trial, word stimuli were individually presented in a sequence of 3 prime words drawn from the same semantic category, and a 4th target word that was either congruent or incongruent with the priming category. All 4 words in a trial sequence were the same length, and they appeared with equal frequency at all 4 presentations within trial sequences across the whole experiment. In the magnitude judgement task, each trial consisted of similarly
sequenced number stimuli that were 6 or 7 numbers in length. In the first 3 priming stimuli, a magnitude change was created by increasing or decreasing the first digit by one as the sequence progressed. The rest of the prime number was randomly generated. At targets, the magnitude change was performed in the 2nd or 3rd digit. There was no change in the first digit between the 3rd prime number and the target. The remaining digits were randomly generated (see Figure 3.2).

3.3.3 Experimental Paradigm

After the EEG electrodes were applied, participants were seated in front of a 21” LCD monitor, with a vertical refresh rate of 60Hz, and viewed trial sequences created using the E-Prime 1.2 software (Psychology Software Tools, Pittsburgh, PA). Stimuli were presented on screen in white, 18-point, centrally-aligned text on a black background. Participants were required to press a button on a keyboard if word targets did not belong to the same semantic category as the primes, or number targets did not carry on the priming magnitude rule (incongruent trials). Congruent trials required no response. The experiment began with a short practise session of 16 trials split equally between word and number sequences. Each trial sequence appeared on screen for 5250 milliseconds: a centrally-aligned white fixation cross presented for 1000 milliseconds. All 3 primes were presented for 500 milliseconds each with a 250-millisecond blank screen between them. The target stimulus was shown on screen for 2000 milliseconds. In total, 200 word trials and 200 number trials were randomly presented in 5 experimental blocks that lasted for 7 minutes each (80 trials per block). In both conditions, 160 trial sequences were congruent, and 40 were incongruent. Participants could rest between blocks.
3.3.4 EEG Data Acquisition

An ActiveTwo Biosemi EEG system (Biosemi Inc., Amsterdam, The Netherlands) with 64 electrodes was used to record continuous brain activity. Electrodes were arranged in an extended 10-20 layout, and fixed to the scalp using an elastic cap studded with plastic electrode holders. All scalp electrode impedances were maintained at less than 20 kOhm. 4 electrodes were placed on the face above and below each eye to record electrooculogram data. Electrode data were digitised with a sampling rate of 512 Hz using ActiView software (Biosemi Inc., Amsterdam, The Netherlands).

3.3.5 Data Pre-processing

EEG Data were pre-processed using the FieldTrip toolbox (version: 20140514; Oostenveld, Fries, Maris, & Schoffelen, 2011) running in MATLAB R2012a (MathWorks, Natick, MA). Data were re-referenced to an average of all 64 electrodes, filtered from 1-60Hz, and visually inspected to remove trials that contained large artefacts resulting from involuntary muscle movements. Independent component analysis (RunICA: Bell and
Sejnowski, 1995) was used to identify artefacts in the data associated with pulse, eye blinks, and lateral eye movements which were discarded. Bad channel data (M = 0.929 channels per participant, SD = 0.997) were repaired from neighbouring channel data using spherical spline interpolation.

3.3.6 Data Analysis
Described briefly, the trial-based TE analysis comprises the following steps: (1) perform source reconstruction on EEG data using LCMV beamformer; (2) extract source time-series across all trials and discretise the data; (3) perform TE analysis across the source time-series using a shifting time-window in all trials; (4) average results of all TE analyses to form group-level TE matrices; (5) use appropriate statistical tests to identify networks of significant information flow within the group.

3.3.6.1 Source time series creation
An LCMV beamformer (Van Veen et al., 1997) was used to create a time series of source activity at each sample point across the entire sequence length in every trial for all participants. A head model was first created using a boundary element model (BEM: Fuchs, Kastner, Wagner, Hawes, & Ebersole, 2002) that was derived from a segmentation of the Colin 27 stereotaxic template (Holmes et al., 1998). Next, a grey matter (GM) restricted source model was created using the following steps: (1) a 20mm resolution regular source grid in MNI template space was created (Figure 3.3a); (2) anatomical labels were defined for each source by interpolating the AAL atlas onto the grid (Tzourio-Mazoyer et al., 2002), and sources not located in grey matter were removed (Figure 3.3b); (3) the GM source grid was inverse-warped to the Colin brain volume (Figure 3.3c); and (4) sources located in cerebellum and deeper structures were removed (Figure 3.3d) because EEG is most sensitive to activity located exclusively in superficial grey matter (Michel et al., 2004). The EEG electrodes were aligned to the scalp shell within the BEM, by transforming a Biosemi electrode layout template using the locations of the nasion, and left and right preauricular points defined on the Colin MRI.
The head model, source model, and electrode positions were used to calculate a leadfield matrix ready for source reconstruction.

Figure 3.3. BEM model and subject source grid: BEM scalp compartment is dark grey, skull compartment is light grey, brain compartment is green, and sources are red. (a) a regular source grid with a resolution of 20 mm is created in MNI-space; (b) MNI-grid is restricted to sources located within GM as defined by the AAL brain atlas; (c) GM-grid is warped to fit the Colin 27 stereotaxic template; (d) 108 sources remain after cerebellar and deep-lying locations are removed.

To obtain source magnitude data from each trial, the EEG trial data were first detrended and averaged before a covariance matrix between 350-450ms after targets was created for each participant. This time-window was chosen because semantic category and number magnitude components were identified between these times in the ERP (Repper-
Day et al., 2016). Using the LCMV method, a spatial filter based on the average across all trials was computed from the covariance matrix and the lead field matrix. Once the averaged spatial filter was obtained, data from each trial was projected through it. The outcome of this process was a source magnitude estimate in 3 dimensions at every grid location for every time point in every trial sequence. We projected the source magnitude to its strongest orientation at every time point, given by the Euclidean norm, and used these values as the source magnitudes for all TE analyses.

3.3.6.2 Discretising the source time series

The first step in calculating the entropy of any time-series is to discretise it into a fixed number of bins. Without discretising the data, the probability distributions $P(X)$ in equation 3.1 would be created from too many outcomes, and TE estimation would become impractical. There are many methods that can be used to discretise the data, but in this analysis we used a direct method (Montemurro et al., 2008). Based on the maximum and minimum of its full range, the signal is separated into bins such that the intervals between each are equal. For example, if we want to separate the signal into 5 bins, and it ranges from -10 and 10, then we divide 20 by 5 to obtain 5 bins of length 4. We used 4 bins to discretise our data because this number has been found to estimate high precision TE using local field potential data (LFP: Lea-Carnall et al., 2011), and both EEG and LFP data are fundamentally similar (Buzsáki et al., 2012). Once discretised, the probability distributions required to calculate the conditional Shannon entropy in equation 3.2 can be estimated by counting the occurrences of each outcome that the sending and receiving signals take (see Figure 3.4 for a pipeline of the direct discretisation method).
Figure 3.4. Direct discretisation method: Calculating the conditional entropy values between 2 signals. (a) original sine wave; (b) after discretisation using 4 bins, the sine wave can take the values 0, 1, 2, or 3; (c) discretised sine wave (blue) is compared to another with a time-delay of 1 (red); (d) the two waves can occupy any of $4 \times 4 = 16$ possible combinations at any one time, and they are counted to estimate the probability distributions used in the entropy equation 3.2.
3.3.6.3 Window TE Analysis

A bespoke trial-based window TE analysis was created to investigate the interactions between source magnitudes across all trials within each participant. Total trials analysed within each participant were equal across the word and number sequences (M = 104.9, SD = 14.1). To assess dynamic changes in connectivity following word and number targets, TE was estimated within a 200ms time-window that was moved across the data beginning when the target was presented. Three windows were analysed: 0-200ms, 200-400ms, and 400-600ms post-target. In equation 3.3, TE is estimated by fixing the sending and receiving data using \( k \) time points in the past of each signal. In practice, it can be difficult to sample the conditioned probabilities because the number of samples required rises exponentially as \( k \) increases (Besserve et al., 2010). Furthermore, the optimum time delays cannot be known before analysis. To overcome these issues, Schreiber (2000) proposed an empirical solution whereby one time delay (\( \tau \)) is used to condition both signals, and estimations of TE are carried out multiple times using a range of delays. In this way, the average of these TE estimations can be used to identify the signals that are most influential in the data (Palus et al., 2001; Honey et al., 2007; Besserve et al., 2010). Instead of estimating TE at time point \( t \) and conditioning on a single time point in the past (\( t-\tau \)), TE was estimated multiple times by conditioning the data at a single time point in the future (\( t+\tau \)) of the receiving signal:

\[
TE_{X \rightarrow Y} = H(Y_{t+\tau}|Y_t) - H(Y_{t+\tau}|Y_t, X_t)
\]  \hspace{1cm} (3.4)

We modified the estimation of TE to incorporate future values of the receiving signal because we are interested in dynamic changes of connectivity immediately after the presentation of our stimuli. If we utilised past values of the sending and receiving signals in our TE analyses, then the conditional probabilities in the 0-200ms window would include pre-target data.
We used a single time delay, which ranged from 10-100ms, to fix the data in 10 separate TE estimations, and the conditional probability distributions were constructed from every data point in the window across all trials (Figure 3.5). Using future values of the receiving signal and a maximum time delay of 100ms in our estimations serves to alter the positioning of the receiving window relative to the sending window during our trial analysis. For example, a sending window of 0-200ms will result in a receiving window of 10-300ms. Mean TE of 10 estimations (10-100ms time delay), was used to identify the sources that were the most active within the window during the experiment. Mean window TE values at each source were arranged into 108x108 matrices of information sent and received between all source pairs to identify the underlying network of source activity within each participant at the window of interest. Finally, the grand average of each participant’s mean window TE matrices following both word and number targets were calculated to ascertain the information flow between sources that were common to the group.

Figure 3.5. **Trial-based window TE analysis between sources**: Inside the time-window, the data points in the receiving signal (red dot), and sending signal (green dot) at time t1, and the data point in the receiving signal at t1 + 50ms (red star) are used to build the conditional probability distributions for the window TE analysis. This is repeated for all data points in n trials. At time t2, the future data point is located outside of the analysis window. Window TE is estimated from the window data from all trials.
3.4 Results

3.4.1 Behavioural data
Data from 2 participants were not analysed due to inadequate task performance and a technical problem with the EEG equipment during recording. Both word (M = 95.2%, SD = 3.4%; t(13) = 50.42, p < 0.001) and number accuracy (M = 89.9%, SD = 3.5%; t(13) = 42.02, p < 0.001) were significantly above 50% chance. However, participants found the number task harder as accuracy was significantly lower relative to words (t(13) = 6.9, p < 0.001), and reaction times following incongruent number trials (M = 1034.3ms, SD = 295.2ms) were significantly slower than words (M = 934.9ms, SD = 294.8ms): t(439) = -5.336, p < 0.001.

3.4.2 TE analysis
After viewing the group level connectivity matrices at all 3 time windows following word targets (see figure 3.6) it was apparent that sources exchanged the least amount of information between 0-200ms, reached a peak at 200-400ms, and reduced down to a level somewhere between the two previous windows by the last. This pattern of information exchange was also seen following number targets (see figure 3.7). The structure of the grand average TE matrices in all windows was dominated by large TE values at occipital sources. Bilateral parietal sources also seem to be sending and receiving relatively large amounts of information regionally and to occipital sources in all windows. There appear to be some sources that are sending and receiving relatively heightened information in frontal and temporal regions, but the patterns of activity are sparse so the 0-200ms matrix exhibits a bias to the bottom right corner. This same bias exists in the 200-400ms connectivity matrix as well; however there does appear to be more wide-ranging activity across the brain. For instance, sources located in frontal regions seem to be sending or receiving more information than they were in the earlier time window (Figure 3.6: middle). Relative to the earliest window, sources in bilateral occipital and parietal lobes seem to be sending and receiving more information as evidenced by larger mean TE values between 200-400ms.
This is especially true for the temporal lobes which appear to be sending to, and receiving from, more bilateral parietal and occipital sources than they were between 0-200ms following word targets. During the 400-600ms window, the pattern of sending and receiving sources seen at 200-400ms is broadly replicated across the whole brain in both hemispheres. Mean TE values appear to have reduced slightly from the middle window, but they are still larger and more widespread than those found earlier in the 0-200ms window. A similar pattern of occipital dominance was found following number targets in all three time windows, and the same spreading of sources to more wide-ranging locations within the later windows could be seen following number targets too (Figure 3.7). There were some subtle differences between the patterns of TE values obtained after both targets. Between 200-400ms and 400-600ms, there appears to be more information exchange on average following numbers as evidenced by the matrices having more of a yellowish hue. TE values seems slightly higher after word targets than numbers, and there appears to be more information being sent to right occipital sources from right temporal areas after word targets between 0-200ms. The maximum TE values in both trials are found between 200-400ms, and both exhibit relatively large amounts of information flow between bilateral occipital and parietal sources. However, word targets seem to produce more information movement to and from temporal sources at 200-400ms relative to numbers.

To compare the gross differences of information exchange between the time windows, mean TE values at each source were collapsed across trial type. Paired samples t-tests revealed that mean source TE was significantly larger between 200-400ms (M = 1.510, SE = 0.007) than it was at 0-200ms (M = 1.435, SE = 0.006): t[3023] = 37.148, p < 0.001. Furthermore, the information exchange between 200-400ms was significantly larger than 400-600ms (M = 1.492, SE = 0.007) too: t[3023] = 26.842, p < 0.001.
Figure 3.6. **Grand average TE Matrices within 200ms windows following word targets**: Grand averages created from all 14 participants’ individual TE matrices at (a) 0-200ms, (b) 200-400ms, and (c) 400-600ms.
Figure 3.7. Grand average TE Matrices within 200ms windows following number targets: Grand averages created from all 14 participants' individual TE matrices at (a) 0-200ms, (b) 200-400ms, and (c) 400-600ms.
Mean source TE between words and numbers were compared at each window to investigate the information exchange across the brain during both tasks (Figure 3.8). No significant difference within the 0-200ms window was found. However, source TE values at 200-400ms following numbers (M = 1.516, SE = 0.009) were larger than those following words (M = 1.504, SE = 0.01) and this difference was significant (t[1511] = 3.331, p = 0.001). Source TE values were also significantly larger for numbers (M = 1.496, SE = 0.009) than they were for words (M = 1.487, SE = 0.009) at 400-600ms (t[1511] = 2.784, p = 0.005).

Figure 3.8. Information exchange at each time-window: The bar chart graphs demonstrate the information exchange (both sent and received) at each source within the window. Information exchange in the middle window was highest, and the differences information exchange following number targets were reliably larger than words between 200-400ms and 400-600ms. Error bars indicate ±2 standard error of the mean source TE.
To investigate differences of intra-regional information flow between word and number trials, mean source TE within and between the left and right hemispheres of the 4 lobes at all 3 time windows (Figure 3.9 and Table 3.1).

**Figure 3.9. Intra-regional TE at each time-window:** Centre matrix shows the location of the regional TE values that were averaged (top-left to bottom-right): frontal, temporal, parietal, and occipital. The bar graphs demonstrate the information exchange (both sending and receiving) within each region for both words and numbers at all time-windows. Error bars indicate ±2 standard error of the mean regional TE.
Table 3.1. Mean intra-regional source TE following both words and numbers in all 3 windows

<table>
<thead>
<tr>
<th>Region</th>
<th>Window</th>
<th>Trial</th>
<th>Source TE (bits/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td></td>
<td>W</td>
<td>1.303</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.348</td>
</tr>
<tr>
<td>Frontal</td>
<td>0-200ms</td>
<td>W</td>
<td>1.352</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.373</td>
</tr>
<tr>
<td></td>
<td>200-400ms</td>
<td>W</td>
<td>1.350</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.369</td>
</tr>
<tr>
<td></td>
<td>400-600ms</td>
<td>W</td>
<td>1.350</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.369</td>
</tr>
<tr>
<td>Temporal</td>
<td>0-200ms</td>
<td>W</td>
<td>1.377</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.337</td>
</tr>
<tr>
<td></td>
<td>200-400ms</td>
<td>W</td>
<td>1.434</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.432</td>
</tr>
<tr>
<td></td>
<td>400-600ms</td>
<td>W</td>
<td>1.395</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.398</td>
</tr>
<tr>
<td>Parietal</td>
<td>0-200ms</td>
<td>W</td>
<td>1.542</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.533</td>
</tr>
<tr>
<td></td>
<td>200-400ms</td>
<td>W</td>
<td>1.598</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.629</td>
</tr>
<tr>
<td></td>
<td>400-600ms</td>
<td>W</td>
<td>1.576</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.600</td>
</tr>
<tr>
<td>Occipital</td>
<td>0-200ms</td>
<td>W</td>
<td>1.724</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.657</td>
</tr>
<tr>
<td></td>
<td>200-400ms</td>
<td>W</td>
<td>1.844</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.808</td>
</tr>
<tr>
<td></td>
<td>400-600ms</td>
<td>W</td>
<td>1.765</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>1.722</td>
</tr>
</tbody>
</table>

Significant differences are in bold.

Paired samples t-tests revealed that intra-regional TE values within frontal sources at 0-200ms were significantly larger for numbers than they were for words ($t[55] = 3.297, p < 0.005$ uncorrected), but despite a trend for larger number TE values in the subsequent time-windows, no further significant differences were found. In temporal sources, word targets produced significantly larger TE values ($t[55] = 2.363, p < 0.05$ uncorrected) between 0-200ms than number targets did. Intra-parietal TE values following number targets were significantly larger than those obtained after word targets between 200-
Finally at occipital sources, intra-regional TE was significantly larger following words relative to number at 0-200ms ($t[55] = 2.217, p < 0.05$ uncorrected), and between 400-600ms ($t[55] = 2.396, p = 0.02$ uncorrected).

We were interested to see whether these regional TE results were affected by the power of the source magnitude signals, as the results of connectivity analyses can be driven by power or signal-to-noise ratio. For example, a positive correlation between TE-estimated information exchange and signal power in more than one frequency band has been found before (Buehlmann and Deco, 2010). Within each participant, the mean source magnitude time series was calculated across bilateral frontal, temporal, parietal, and occipital sources at each time-window following both words and numbers. Next, the power of these regional source signals was calculated as the square of its root-mean-square level. Finally, the grand mean regional source power within all three time-windows following both words and numbers was calculated (Figure 3.10).

Paired-samples t-tests revealed that none of the regional differences between mean source signal power following words and numbers were significant. The largest regional signal powers and TE values were both found at occipital sources. However, this is about the only similarity between the two analyses. Even the consistent patterns of information exchange between frontal sources, and occipital sources, at each time-window were not repeated in the power analysis. This comparison of regional TE and power analyses seems to indicate that there is more than a simple high/low power equals high/low TE explanation for the patterns of information flow found using our analysis method.
To investigate task-dependent information exchange across the brain, mean sending and receiving source TE were calculated between all regions at all 3 windows following both words and numbers. Particular focus was placed on the temporal and parietal lobes, as they are thought to be especially important for processing semantic category and number magnitude respectively. Temporal sources (Figure 3.11) were sending significantly more information ($t[55] = 2.425$, $p = 0.034$ uncorrected, $r = 0.282$) to occipital sources following words ($M = 1.529$, $SE = 0.052$) than numbers ($M = 1.483$, $SE = 0.041$)
between 0-200ms. Before we ran this analysis, we hypothesised that we would see a convergence of neural activity from occipital regions to anterior temporal lobe sources around 400ms after word targets were presented. We did find higher TE values from occipital sources to temporal sources for words relative to numbers between 200-400ms, but the difference did not quite reach significance (p = 0.088). Temporal sources were receiving significantly more information \( (t[55] = 2.423, p = 0.019 \text{ uncorrected}, r = 0.311) \) from frontal sources following number targets \( (M = 1.336, SE = 0.041) \) than they were following words \( (M = 1.305, SE = 0.0423) \) between 0-200ms. This same pattern of information flow was repeated at 200-400ms and again it was significant: numbers \( (M = 1.40, SE = 0.042) \); words \( (M = 1.354, SE = 0.41) \); \( t[55] = 3.687, p = 0.001 \) uncorrected, \( r = 0.445 \). Interestingly, temporal sources started sending significantly more information back to frontal sources \( (t[55] = 2.425, p = 0.019 \text{ uncorrected}, r = 0.311) \) following numbers \( (M = 1.407, SE = 0.045) \) than they did after words \( (M = 1.375, SE = 0.044) \) between 200-400ms. According to these results, frontal and temporal sources are exchanging information at a significantly higher rate for numbers than they did for words. This enhanced interaction following numbers seems to start at frontal sources between 0-200ms before becoming 2-way communication between the regions 200-400ms after presentation of targets.

At parietal sources (Figure 3.12) between 0-200ms, we found that more information was being received from occipital sources following word targets \( (M = 1.639, SE = 0.049) \) than for numbers \( (M = 1.596, SE = 0.043) \). This difference was significant \( (t[55] = 2.388, p = 0.02 \text{ uncorrected}, r = 0.307) \). In the parietal to frontal direction, the difference between numbers \( (M = 1.491, SE = 0.04) \) and words \( (M = 1.458, SE = 0.042) \) was also significant \( (t[55] = 3.457, p < 0.001 \text{ uncorrected}, r = 0.423) \), and the same was true in the opposite direction: numbers \( (M = 1.449, SE = 0.39) \); words \( (M = 1.414, SE = 0.41) \); \( t[55] = 3.427, p < 0.001 \text{ uncorrected}, r = 0.42 \). So it would appear that an increased 2-way exchange of information between frontal and parietal sources is occurring following numbers, relative to words, between 200-400ms.
Figure 3.11. Information flow between temporal sources and the rest of the brain: Bar graphs depict information outflow from temporal (top) sources to the rest of the brain, and inflow from the rest of the brain to temporal sources (bottom) after both words and numbers at all 3 time-windows.
Figure 3.12. Information flow between parietal sources and the rest of the brain: Bar graphs depict information outflow from parietal (top) sources to the rest of the brain, and inflow from the rest of the brain to parietal sources (bottom) after both words and numbers at all 3 time-windows.
3.5 Discussion

Our TE analysis confirmed that the largest amounts of information in both word and number group average TE matrices were being exchanged within and between bilateral occipital at all time windows. The presentation of visual task stimuli illicit a common sequence of C1, P1, and N1 components in the event-related potential that are collectively referred to as the visual-evoked potential (VEP) (Jeffreys and Axford, 1972). The cortical generators of the C1 component have been localised to primary visual cortex, the early phase of P1 has been attributed to middle occipital gyrus, and the late phase of P1 and posterior N1 components are thought to arise in the ventral extrastriate cortex of the fusiform gyrus (Di Russo et al., 2002). Significant information flow was found at sources in many of our participants at these locations using our analysis method. Furthermore, the patterns of TE seemed to shift from being relatively focussed at occipital sources towards more wide-spread activity as the trial progressed. This more wide-spread activation underpinned a clear pattern of overall information movement between the time-windows: both word and number targets elicited the lowest amount of information exchange between 0-200ms, rose to a peak between 200-400ms, and settled somewhere in the between at 400-600ms. This pattern could be representative of the shift from local automatically-evoked processes which extract information about stimuli to more sophisticated cognitive processes that arise, as a response to stimuli, from interactions between multiple neural populations. This cascade of neural processing is represented in the classic split between exogenous and endogenous components of the ERP. The early exogenous components are automatic responses whose characteristics are determined by the physical properties of the stimuli, and they can be seen in the first ~250ms of the ERP (Donchin et al., 1978). Whereas endogenous components are associated with cognitive processes, such as decision making, which arise in response to a stimulus and are moulded by the task being performed (Picton et al., 2000). The endogenous components are visible in the ERP between ~250-1000ms after stimulus
presentation, therefore the reliable increase of TE values that we have found in the 200-400ms and 400-600ms fit well with these time-frames.

3.5.1 Information flow following word targets
We hypothesised that more information would flow from occipital sources towards temporal regions following word targets relative to numbers at about 400ms. This prediction was based on studies that highlighted the importance of the temporal lobes for semantic processing, and the N400 ERP component which is modulated by semantic effects (Marinkovic et al., 2003; Patterson et al., 2007; Pobric et al., 2007; Visser and Lambon Ralph, 2011; Jackson et al., 2015). Within temporal regions we found that information exchange was significantly larger for words relative to numbers between 0-200ms, but there was no real difference between the trial types during the later windows. When we looked at the directions of information flow between brain regions we found that temporal sources were sending significantly more information to occipital sources at 0-200ms, and the return direction (from occipital to temporal) almost reached significance between 200-400ms. Indeed, there seemed to be a preferential 2-way interaction between occipital and temporal sources for words at 0-200ms and 200-400ms as mean TE values were consistently higher. Traditional models of semantic processing posit that stimuli features are extracted at primary sensory regions, and that this information converges at the anterior temporal lobes to form semantic concepts at about 400ms (Marinkovic et al., 2003; Patterson et al., 2007; Jackson et al., 2015). The results of our analysis seem to align with a model that suggests there are 2 stages of semantic processing at the temporal lobe (Bar, 2006). In this model, some semantic processing occurs early in the temporal lobes and the results are sent back to more posterior brain regions to inform feature extraction and the subsequent concept formation. Evidence for this early top-down semantic processing has been found using implanted electrodes which tracked semantic information from anterior temporal lobes towards posterior regions as early as 130ms after stimuli were presented (Chan et al., 2011).
An interesting result using this analysis was found within occipital sources: words resulted in larger TE values at all 3 time-windows relative to numbers. Indeed, between 0-200ms and 400-600ms these differences were significant. The result at 0-200ms could be linked to the reliably larger information flows from temporal to occipital, and from occipital to parietal that were found in the same window. The latter interaction could reflect the processing of semantic word properties that are supported by both occipital and parietal sources around 200ms after visually presented words (Hauk et al., 2006). Both temporal and parietal sources are consistently receiving more information from occipital sources following words in the first 2 time-windows (see Figures 3.11 and 3.12; bottom graphs). Visual recognition processes of words and numbers share common neural substrates, but there is some evidence to suggest that there are separate systems for both stimulus types (Allison et al., 1994). Once a word is recognised, encoding of the object’s name, size, shape, motion, colour, environment, texture, usage, etc. will occur (Patterson et al., 2007); whereas the presentation of number stimuli will most likely invoke activity related to the encoding of number quantities such as size, even/odd, place-value (Grossberg and Repin, 2003). Indeed, there is some evidence to suggest that the parietal lobe, rather than the occipital, is the first area that extracts visual numerical information in primates (Nieder and Miller, 2003). Therefore, word stimuli describing objects could potentially invoke more activity than numbers at occipital locations dedicated to object feature extraction.

3.5.2 Information flow during the number magnitude judgment task
Total information exchange across all sources was the same for numbers and words at 0-200ms, but significantly higher for numbers at 200-400ms and 400-600ms windows (see figure 3.7). This result was probably driven by a difference in working memory load between the two tasks: it is harder to remember the magnitude of the 3rd number stimuli (a 6 or 7-digit number), than it is to remember the semantic category of the preceding 3 word stimuli. Behavioural data indicated that participants found the numbers task significantly more difficult than the words: reaction times were longer and participants
were less accurate on the numbers task. When we looked for information exchange
differences between words and numbers within frontal sources, we found that TE values
following numbers were higher than those following words at all 3 time-windows, and
significantly so at 0-200ms. Interestingly, this finding could be driven by regional
information outflow from frontal to temporal sources, which was reliably higher following
numbers between 0-200ms. By 200-400ms, significantly increased 2-way information
exchange was occurring between frontal and temporal sources following numbers.
Working memory processes are commonly associated with activity in dorsolateral
prefrontal cortex (DLPFC) (Owen, 1997; Curtis and D’Esposito, 2003). Furthermore,
increasing working memory load has been found to produce a proportional rise of theta
and reduction of alpha within frontal cortex, which would suggest that more cortical
resources are expended during harder WM tasks (Gevins et al., 1997). The medial
temporal lobe (MTL) has also been implicated in WM tasks (Ranganath and D’Esposito,
2001; Nichols et al., 2006; Olson et al., 2006), and working memory load has also been
found to affect neural activity in the medial temporal lobe (Axmacher et al., 2007). We
propose that the TE analysis highlighted the interactions between DLPFC and MTL
underpinning working memory processes. If this is the case then it appears that working
memory processes start early (0-200ms) in frontal regions, before recruiting areas in
temporal lobe. This result could add more support to the early top-down facilitation
model of visual object recognition (Bar, 2006).

We hypothesised that we would find information flow towards the parietal lobe following
numbers targets at about 200-400ms. This prediction was driven by the extensive
literature highlighting the importance of parietal regions to number processing (Dehaene
et al., 2003, 2004; Nieder and Dehaene, 2009), and our own ERP analysis of this data
which found a component linked to magnitude judgement processes about 365ms after
number presentations (Repper-Day et al., 2016). Information exchange between sources
in both parietal hemispheres was reliably larger for numbers between 200-400ms. This
result was driven by the interaction between frontal and parietal sources that was
significantly larger following numbers than words. This information exchange started from frontal to parietal sources between 0-200ms and became a 2-way interaction by 200-400ms. This could represent some facet of working memory aiding the number magnitude judgement, as WM task-related activity is often found in both frontal and parietal cortices (Curtis, 2006; Moore et al., 2006; Klingberg, 2010). An increase in connectivity has been found between the 2 regions as WM load increases (Honey et al., 2002); WM training has been found to increase activity in frontal and parietal regions (Olesen et al., 2004); and superior parietal cortex is critical for the kinds of WM manipulations that are required for accuracy in our numbers task (Koenigs et al., 2009).

3.5.3 Future Work

These results, although encouraging, must be tempered with the knowledge that there is much refinement work that needs to be carried out to establish confidence in the method. Firstly, it appears that we might have needed to recruit more participants. We did find some reliable regional information exchange differences between words and numbers at certain time-windows, but some predictable information movements, such as more information following words being sent to temporal regions from occipital sources between 200-400ms, didn’t quite reach significance. This might have been caused by variability of the TE levels between participants. Furthermore, although the occipital activation is a good result, there is no subtlety to the patterns within the region: most of these occipital sources seem to be “on” during the time-windows. It is reasonable to suggest that occipital sources should be active during visual task, but the fact that we calculate mean TE over a range of time delays and our relatively large time-windows probably contribute to a blurring of the true network activity during each window. Future work needs to be carried out on the optimisation of time-delays and assessing how small we can make the window without biasing the TE calculation. Furthermore, our code needs to be optimised to estimate TE more efficiently so we can increase the number of sources in our analysis to improve its spatial resolution. Only then might we gain a
richer picture of the transient short- and long-distance neural activity that underpins cognitive performance on these tasks.

3.5.4 Conclusion

Our goal was to ascertain whether transfer entropy could be used to investigate causal interactions within EEG source data. We have found evidence of high levels of information flow emanating from occipital areas within our task data, not surprising as our experimental stimuli were visual. We also found a plausible pattern of total information flow within our time-windows that was consistent with the idea that early task processing is mainly dedicated to feature extraction in sensory regions, before progressing to higher level cognitive processing shared between regions. Moreover, we found reliably higher information exchange between regions at various time windows for both words and numbers that mirrored previous semantic and number processing research.

A method that can be used in vivo to ascertain the real-time direction and timing of regional interactions that underpin brain function would be a valuable addition to the battery of connectivity analysis used in neuroscientific research. We believe that the results presented here suggest that, with more refinement, this trial-based TE analysis of source data could be that method.
Chapter 4

Tracking the effects of TMS using a novel trial-based Transfer Entropy analysis

Chris Repper-Day, University of Manchester, UK
Gorana Pobric, University of Manchester, UK
Stephen R. Welbourne, University of Manchester, UK
Marcelo A. Montemurro, University of Manchester, UK
4.1 Abstract

Connectivity can be investigated by tracking neural activity evoked by transcranial magnetic stimulation (TMS) from a target site to the rest of the brain. We developed a novel connectivity analysis using Transfer Entropy, and applied it to source reconstructed TMS-EEG data. TE is a data-driven quantifier of connectivity that can infer the direction of linear and non-linear information exchange between signals over a range of time-delays. Online TMS was applied to bilateral ATL and IPS, and the analysis technique was able to map the influence between 0-200ms, 200-400ms, and 400-600ms post-TMS. The ATL and IPS were chosen as they are thought to be important hubs within the networks that support semantic and number processing respectively. Information flow across the entire brain peaked in the 200ms immediately following the TMS pulse and returned towards baseline in the subsequent time-windows. At 0-200ms, patterns of left ATL TMS influence were located primarily in wide-spread left temporal lobe, and inferior frontal and inferior occipital cortices, and exhibited spatial similarities to the results of neuroimaging experiments that probed semantic processing. The analysis of right ATL TMS obtained a mirror image of the left. Left parietal stimulation resulted in a bilateral parietal, superior occipital, and superior prefrontal influence, which extended slightly further in the ipsilateral hemisphere to stimulation site. Again, the results at the contralateral site were a virtual mirror image.
4.2 Introduction

To fully comprehend how the brain works we must consider the dual principles of *functional segregation* and *functional integration*. Many neuroimaging investigations have utilised functional segregation to identify areas of the brain whose activities are statistically linked to experimental tasks. However, the importance of functional integration has only recently been highlighted by influential theories which posit that brain function arises from the synchronous interactions between the nodes of connected brain networks (Engel et al., 2001; Varela et al., 2001; Buzsáki, 2006). Cognitive neuroscience has adopted an analytical framework to describe the types of connectivity between brain regions, and functional integration is mediated by 2 of these: *functional* and *effective connectivity* (Friston, 1994). Functional connectivity (FC) is defined as a statistical dependency between activities at distinct neuronal populations (Friston, 2011), and essentially describes a pattern of synchronised neuroimaging signal dynamics (Lee et al., 2003). Effective connectivity (EC) describes the influence that these brain areas have over each other (Friston, 2011). This influence can be direct or indirect, and is dependent on the anatomical connections between the functionally connected brain areas (Friston, 1994).

4.2.1 Neuroimaging of connectivity

Connectivity has been investigated with many neuroimaging techniques including fMRI, PET, TMS, EEG, and MEG. However, whilst fMRI and PET have excellent spatial resolution, they are too slow to quantify the changes in connectivity that occurs within functional brain networks. Conversely, TMS possesses excellent temporal resolution and is able to influence activity within brain networks during an experimental task at millisecond time-scales. TMS investigations of connectivity have utilised a dual-pulse paradigm to show that induced activity at a target area can influence responses at functionally connected regions. Dual-pulse experiments consist of a target pulse being delivered first to motor hand region, and a second conditioning pulse being delivered at a remote, but functionally connected site. The influence of the conditioning pulse is then
assessed on the motor-evoked potential (MEP) generated by the original target pulse (Angstmann and Siebner, 2012). Using the dual-pulse method, M1Hand has been shown to be functionally connected with hand homologue region (Ferbert et al., 1992), supplementary motor area (Civardi et al., 2001), dorsal premotor cortex (Bäumer et al., 2006), and posterior parietal cortex (Koch et al., 2007). However, the dual-pulse paradigm is limited to studying areas of the brain that produce a marker of excitability such as the motor (MEP) or visual (phosphenes) cortices. Unlike fMRI and PET, both EEG and MEG are direct measures of brain activity with temporal resolutions that are able to capture the transient interactions between brain regions, and as such they are optimal for studying functional and effective connectivity (Gross et al., 2001).

4.2.2 Quantifying connectivity within M/EEG data

Many quantifiers of FC and EC have been applied to M/EEG data (see Greenblatt, Pflieger, & Ossadtchi, 2012; Pereda, Quiroga, & Bhattacharya, 2005 for extensive reviews), however they have mostly been applied to sensor data which is far from ideal because the spreading of the fields can confound results. For example, the FC measure known as coherence (which is the cross-correlation between signals in the frequency domain) can be artificially inflated in nearby sensors due to overlapping fields (Greenblatt et al., 2012). Recently, some researchers have argued that connectivity investigations utilising M/EEG should be investigated at the source level to circumvent this problem (Schoffelen and Gross, 2009). One such technique that assesses source-level connectivity is dynamic imaging of coherent sources (DICS: Gross et al., 2001). DICS identifies brain networks by using a spatial filter algorithm to identify brain sources that possess coherence at a particular frequency band. The spatial filter can be applied to sources located throughout the whole brain, and correlations with brain sources exhibiting maximum frequency power can be used to find networks of functionally connected sources (Gross et al., 2001). DICS has been used to show that controlled finger movements correlated with oscillatory activity between 6-9Hz in a cerebello-thalamo-cortical network (Gross et al., 2002), and it was also able to detect the behaviour of brain networks that contributed to
different types of epileptic seizure (Moeller et al., 2013). However, because DICS utilises coherence (a quantifier of FC), it is insensitive to causality within the networks it defines. Indeed, Moeller and colleagues (2013) used partial directed coherence to investigate EC between their sources of interest. Additionally, interpretation of EC results can be difficult because interdependence between activities observed in two neural populations may or may not be due to a causal influence that one exerts over another. There could be other regions mediating activity at both (Lee et al., 2003). Furthermore, a significant amount of brain activity is internally generated, and is known to affect how neuronal populations respond to stimuli (Arieli et al., 1996). A less ambiguous method of investigating effective connectivity would be to perturb the brain, using TMS, and observe how the disruption influenced activity at inter-connected brain regions.

4.2.3 Investigating connectivity using combined TMS-EEG

Combined TMS-EEG has been suggested as an excellent method of investigating brain function as it can provide a real-time picture of cortical reactivity and connectivity. TMS has been found to produce a consistent pattern of deflections in EEG data, known as the TMS-evoked potential (TEP), at many areas of the brain including those that do not produce an overt physical marker of excitability (Ilmoniemi and Kicić, 2010; Miniussi and Thut, 2010). For instance, TMS was applied to the same target in prefrontal cortex at 4 different intensities, and the resulting TEPs exhibited a consistent pattern of deflections and scalp topographies which indicated that the same cortical network was activated. Furthermore, the amplitude of the deflections were positively correlated with stimulation intensity (Kähkönen et al., 2005). TMS applied to right frontal eye field affected neural activity evoked by visual stimuli at posterior electrodes in the same hemisphere, as well as anticipatory activity before stimuli presentations (Taylor et al., 2007). Furthermore, TMS-EEG has shown that cortical connectivity is reduced during sleep (Massimini et al., 2005). The suitability of TMS-EEG to connectivity research is encouraging; however the problem of field spread can cause difficulties in interpretation of results. Finding TEP effects at distant electrodes does not tell us too much about the location of the actual
brain areas that are connected to the target site because scalp topographies can be produced by multiple configurations of internal sources and not necessarily by those located below the sensors (Michel et al., 2004). Therefore, the potential advantages to connectivity investigations that utilise TMS-EEG are undermined when the combined data is analysed at the channel level. Theoretically, an analysis of TMS-EEG source–level data using a quantifier of EC could elucidate the real-time dynamic influence of TMS within a network of brain regions.

4.2.4 Aims of the current study

Inspired by the DICS method (Gross et al., 2001), we created a novel analysis that looked for patterns of connectivity between EEG source data using the quantifier of EC known as transfer entropy. Transfer entropy (TE) is a non-parametric statistic that can measure how much directed information is sent between signals. It has recently been championed as a useful addition to the battery of electrophysiological connectivity quantifiers (Vicente et al., 2011). Theoretically, TE should be able to map the influence of a TMS pulse between EEG source data at millisecond time scales, and, by extension, the causal interactions that arise within brain networks after TMS-evoked activity. We collected continuous EEG data following TMS application to bilateral anterior temporal lobes (ATLs) and bilateral superior parietal lobes around the intraparietal sulcus (IPS). We chose these locations as they are important nodes in the widely distributed brain networks that support the semantic (Patterson et al., 2007) and number processing (Dehaene et al., 2003) systems respectively. The effects of TMS at a target site can be both inhibitory or excitatory, and are modulated by factors including intensity and coil angle (Siebner et al., 2009). It is therefore difficult to predict how the TMS effect will manifest itself in a map of EC. TMS-evoked activity at the target could lead to an increase of information transfer towards connected regions, or a reduction at the target and its connected regions could be observed. Alternatively, increased activity at the target could disrupt synchronous activity between connected areas leading to a reduction of information sent and received.
4.3 Method

4.3.1 Participants
6 volunteers (4 female, 2 male; mean age = 21.66 [SD = 2.34] years) were recruited to take part in the study. All participants were right-handed, had no history of neurological disorders, and possessed normal or corrected-to-normal vision. All participants gave informed consent, and the experiment was covered by the research group's ethical approval.

4.3.2 Experimental Paradigm
Participants had EEG electrodes applied, were seated in front of a 21” LCD monitor with a vertical refresh rate of 60Hz, and requested to relax and watch a silent movie presented on screen. The movie was “Journey to the edge of the universe”, which was chosen because it contained no arousing or emotional imagery, and no human faces or items of personal significance. The movie was incorporated into the paradigm in an attempt to focus the participants’ attention away from the stimulation and reduce boredom. Whilst viewing the movie, participants were subjected to TMS during 8 experimental blocks. During the first 4 blocks, the TMS location was alternated between locations in left and right anterior temporal lobes, and left and right superior parietal lobes. Locations were counter-balanced across participants such that 2 sequential stimulation blocks never occurred in the same lobe. The 2nd and 3rd blocks always occurred in the same hemisphere, so their locations were also counter-balanced across participants: parietal first or temporal first. After the first 4 blocks, the location sequence was repeated. Each block lasted approximately 2 minutes 45 seconds, and there were 2-minute rest periods between blocks. The entire experiment lasted approximately 36 minutes.
4.3.3 EEG data acquisition

Continuous brain activity was recorded at 64 electrodes, in an extended 10-20 layout, using an ActiveTwo Biosemi EEG system (Biosemi Inc., Amsterdam, The Netherlands). The electrodes were fixed to the scalp using an elastic cap studded with plastic holders. All electrode impedances were maintained at less than 20 kOhm. The Biosemi system was connected to a computer running ActiView software (Biosemi Inc., Amsterdam, The Netherlands), and the data were digitised with a sampling rate of 2048 Hz.

4.3.4 TMS Paradigm

TMS was delivered online, during the concurrent visual task, using a MagStim Rapid2 (MagStim Co., Whitland, UK) stimulator with two external boosters (maximum output = 2.2 Tesla approx.) fitted with a 70mm figure-of-eight coil. Motor threshold (MT) was determined for each participant after the EEG cap was fitted: mean 75% maximum stimulator output, SD = 5.4%. During each stimulation block, single pulse TMS was delivered with an inter-pulse interval of 1 second plus a random jitter between 0-0.2 seconds. TMS pulses were coordinated between the stimulator and the EEG recording software using E-Prime 1.2 software (Psychology Software Tools, Pittsburgh, PA). Left and right ATL stimulation was administered 2cm below the mid-point between electrodes T7 and FT7, and T8 and FT8 respectively. Left and right parietal stimulations were delivered at the mid-point between CP1, CP3, P1, and P3, and CP2, CP4, P4, and P2 respectively (see Figure 4.1). P3 and P4 are known to be located above IPS activation found using fMRI (Klein et al., 2013).
Figure 4.1. Sites of both temporal and parietal TMS: (left) location of left ATL stimulation is shown in red. Stimulation was delivered 2cm below the mid-point of electrodes FT7 and T7 which are coloured blue. Process was repeated in the other hemisphere between electrodes FT8 and T8 for right ATL TMS; (right) location of left IPS stimulation is shown in red. Stimulation was delivered between the 4 blue electrodes: CP1, CP3, P1, and P3. The location of right IPS stimulation site can be seen within the glass brain as well.

4.3.5 Data pre-processing

EEG Data were pre-processed and analysed in MATLAB R2012a (MathWorks, Natick, MA) using the FieldTrip toolbox (version: 20140514: Oostenveld et al., 2011). Data were re-referenced to an average of all 64 electrodes, and visually inspected to remove trials that contained large artefacts resulting from involuntary muscle movements. Next, TMS-related artefacts were removed using peer-reviewed procedures (Herring et al., 2015) detailed in a tutorial on the FieldTrip website (www.fieldtriptoolbox.org/tutorial/tms-eeg). First, the step response of the EEG amplifier results in a large “ringing” artefact that could be seen within the first 10ms after a pulse (See Figure 4.2a). This portion of the signal was removed to produce pre- and post- ringing sections, and independent component analysis (RunICA: Bell and Sejnowski, 1995) was performed on the segmented data.
Components were removed from the data that represented the large exponential decay artefacts (Figure 4.2b) which are thought to be caused by muscle contractions/moving electrodes (Ilmoniemi & Kicić, 2010); residual muscle artefacts (Figure 4.2c); and eye blinks and lateral eye movements (Figure 4.2d). We identified any remaining components that captured 50Hz line noise by calculating power spectrum densities, and these were also removed (Figure 4.2e). Following component rejection and back projection of the data, it was found that the TMS-related muscle artefacts could not be fully removed by ICA during the first 27ms period. Therefore, this section was excised from the data and cubic interpolation was used to fill the gap between the pre- and post-pulse segments (Figure 4.2f). Next, the cleaned data were subjected to Butterworth low-pass filter (cut-off frequency of 200Hz) to avoid aliasing. Mean and SD were calculated over all times and trials, the data were z-transformed, and any trials deviating by 5 SDs were removed. At this point any trials containing residual muscle artefacts were removed following a last visual inspection. Finally, the effects of slow fluctuations to the variance of faster components were reduced using a bandpass filter with cut-off frequencies between 3 and 45 Hz. After the entire cleaning pipeline, we were left with an average of 177.33 (SD = 9.22) trials per stimulation location.
Figure 4.2. TMS-EEG data cleaning pipeline: All images are derived from a subject who had TMS applied to right ATL. (a) Removal of step response by excising the first 10ms of data after TMS pulse. T8 is one of the electrodes closest to the stimulation site; (b) Time-course and scalp topographies of 2 decay components. The TMS pulse can be seen in right temporal and anterior frontal electrodes; (c) Muscle components: the red component is TMS-evoked, but the blue component has captured a non-TMS muscle artefact; (d) Eye components: the red time component is blinking, and the blue captures lateral eye movements; (e) Line noise; (f) Time course of cleaned data at electrode Cz, showing a typical looking TMS-evoked potential.
4.3.6 Data analysis

We created a time series of source activity at each sample point across the entire sequence length in every trial for all participants using an LCMV beamformer (Van Veen et al., 1997). Connectivity analyses applied to LCMV-reconstructed simulated sources have produced better results than other popular techniques like minimum norm estimation (Schoffelen and Gross, 2009). The head model was created using a boundary element model (BEM: Fuchs, Kastner, Wagner, Hawes, & Ebersole, 2002) that was derived from a segmentation of the Colin 27 stereotaxic template (Holmes et al., 1998). Next, a grey matter (GM) restricted source model was created using the following steps: (1) a 20mm resolution regular source grid in MNI template space was created; (2) anatomical labels were defined for each source by interpolating the AAL atlas onto the grid (Tzourio-Mazoyer et al., 2002), and sources not located in grey matter were removed; (3) the GM source grid was inverse-warped to the Colin brain volume; and (4) sources located in cerebellum and deeper structures were removed because EEG is most sensitive to activity located exclusively in superficial grey matter (Michel et al., 2004). The EEG electrodes were aligned to the scalp shell within the BEM, by transforming a Biosemi electrode layout template using the locations of the nasion, and left and right preauricular points defined on the Colin MRI. The head model, source model, and electrode positions were used to calculate a leadfield matrix ready for source reconstruction.

A bespoke trial-based TE analysis was created to investigate the interactions between source magnitudes across all trials within each participant. TE measures the directed transfer of information between two signals using the Shannon entropy $H$ (Shannon, 1948), which is the cornerstone of information theory. The Shannon entropy quantifies the information contained in a signal by measuring uncertainty. A decrease of uncertainty leads to an increase of information. The information in a system $X$ can be determined by calculating the entropy of the probability distribution of all its values.
The Shannon Entropy is defined as the average uncertainty of the values $x$ weighted by the probability of their occurrence $p(x)$. If $X$ contains $M$ values, uncertainty $H(X)$ now becomes:

$$H(X) = - \sum_{x \in M} P(x) \log_2 P(x)$$  \hspace{1cm} (4.1)

Transfer entropy from signal $X$ to signal $Y$ is the increase of information gained about values of $Y$ using the past of $X$ and $Y$. $TE$ is calculated using equation 4.2. $H(Y_t \mid Y_t^{(k)})$ represents the Shannon entropy in signal $Y$ at time $t$ when it has been conditioned (or fixed) to its own past $k$ time points. $H(Y_t \mid Y_t^{(k)}, X_t^{(k)})$ represents the Shannon entropy of $Y$ at time $t$ when it has been conditioned to $k$ past values of $X$ as well as $k$ past values of $Y$. If the incorporation of past values of $X$ increases the information within $Y$ more than using the past of $Y$ alone, then $X$ is said to cause, or influence, $Y$.

$$TE_{X \rightarrow Y} = H(Y_t \mid Y_t^{(k)}) - H(Y_t \mid Y_t^{(k)}, X_t^{(k)})$$  \hspace{1cm} (4.2)

In practice, it can be difficult to sample the conditioned probabilities because as $k$ increases, the number of samples required rises exponentially (Besserve et al., 2010). Furthermore, the optimum time delays cannot be known before analysis. As a solution, TE can be estimated repeatedly using a range of time delays, and the mean of these estimations can be used to identify the effect of causation (Schreiber, 2000). The conditional probability distributions required for TE estimation were constructed from every data point in the window across all trials.
Instead of estimating TE at time point t and conditioning on a single time point in the past t-\(\tau\), TE was estimated multiple times by conditioning the data at a single time point in the future of the receiving signal:

\[
TE_{X\rightarrow Y} = H(Y_{t+\tau}|Y_t) - H(Y_{t+\tau}|Y_t, X_t)
\]  

(4.3)

TE was estimated within a 200ms time-window that was moved across the data beginning 300ms before the TMS pulse and ending 600ms after. We incorporated future values of the receiving signal because we are interested in dynamic changes of connectivity immediately after the TMS pulse. If we utilised past values of the sending and receiving signals in our TE analyses, then the conditional probabilities in the 0-200ms window would include pre-TMS data. Using future values of the receiving signal and a maximum time delay of 100ms in our estimations serves to alter the positioning of the receiving window relative to the sending window during our trial analysis. For example, a sending window of 0-200ms will result in a receiving window of 10-300ms. TE was estimated 10 times between sending and receiving sources at every data-point within the window (\(t\)) using a range of 10 time delays (\(\tau\)) from 10-100ms. As in equation 4.3, the source magnitude at the sending source (\(X_t\)), the source magnitude at the receiving source (\(Y_t\)), and the source magnitude at some point in the future of receiving source (\(Y_{t+\tau}\)) were used to build up a distribution of source magnitudes from which TE could be estimated. This was repeated for each time delay, in every 200ms analysis window, across all trials (Figure 4.3). We used multiple time delays because the optimum cannot be known before analysis (Besserve et al., 2010). Mean TE using the 10-100ms range of \(\tau\) was used to identify the sources that were the most active within the window after TMS.
Figure 4.3. Trial-based window TE analysis between sources: The contributions of points $t_1$ and $t_2$ to TE estimation ($\tau = 50$ms) within a 200ms window are shown. Window TE is calculated as the mean of TE estimated from $n$ trial windows. Because $t_2$ is towards the end of the window, the receiving channel TE can be estimated at a time-point that exceeds the end time-point in the window.

Four windows were analysed: -300--100ms pre-TMS, and 0-200ms, 200-400ms, and 400-600ms post-TMS. The baseline window was chosen so that there was no contaminating effect of pulse data in the receiving window. Window TE values at each source were arranged into 108x108 matrices of information sent and received between all source pairs to identify the underlying network of source activity within each participant at the window of interest. Finally, the grand mean of all participants’ window matrices were combined to identify information flow that was consistent across the group following TMS.
4.4 Results

4.4.1 TMS evoked EEG activity

Grand average ERP calculated from all participants and all stimulation sites revealed that TMS-evoked activity could be seen within electrode data at wide-spread locations across the scalp (Figure 4.4).

![Figure 4.4. Grand average of TMS-evoked activity](image)

**Figure 4.4. Grand average of TMS-evoked activity**: Grand average ERPs calculated from data recorded after stimulation to all 4 target sites are shown at each electrode location. An example of the TEP waveform is presented from FCz. Each TMS pulse occurred at time 0ms.
This was manifested in a characteristic sequence of deflections in the ERP that resemble TEP from other studies. The largest deflections can be seen between 100-250ms following TMS at fronto-central electrodes along the midline such as FCz and Cz. Furthermore, it appears that the deflections in electrode data located near to stimulation sites (lATL = F7, FT7; rATL = F8, FT8; Parietal = P3, P1, Pz, P2, P4) are much smaller than those seen elsewhere. This effect could be caused by the removal of genuine TMS-evoked activity along with muscle artefacts, which were most pronounced at these locations.

4.4.2 Comparison of information flow pre- and post-TMS across the whole brain

The grand average window matrices before TMS at all 4 sites revealed that the largest amounts of information were typically being exchanged between occipital sources (see the top left matrix of Figures 4.5, 4.6, 4.7, and 4.8). This is not an unreasonable result because our participants were simply processing the moving images that were being presented during this time-window. Immediately following TMS to left ATL we can see an obvious increase of TE values between source pairs located throughout the brain (see the top right matrix of Figures 4.5, 4.6, 4.7, and 4.8). Furthermore, TE values were larger during the 0-200ms post-TMS window than at all the other windows. In effect, it appears that TMS has evoked a transient increase of information flow around the brain in both hemispheres. Interestingly, a consistent increase of TE between the sources located within the brain region that was targeted by TMS can be seen: it is particularly stark in the sources located in right temporal and right parietal lobes after stimulation in the vicinity (compare the top 2 matrices in figures 4.6 and 4.8). Following this initial increase, TE values across the brain appear to reduce towards the levels found pre-TMS.
Figure 4.5. Grand average source TE matrices pre- and post-TMS to IATL: Sending and receiving TE between all source pairs after TMS was applied to left anterior temporal lobe during pre-TMS baseline (top left), 0-200ms (top right), 200-400ms (bottom left), and 400-600ms (bottom right).
Figure 4.6. Grand average source TE matrices pre- and post-TMS to rATL: Sending and receiving TE between all source pairs after TMS was applied to right anterior temporal lobe during pre-TMS (top left), 0-200ms (top right), 200-400ms (bottom left), and 400-600ms (bottom right). A relatively large increase in TE, from pre- to post-TMS, can be seen in right temporal sources close to the stimulation site between 0-200ms.
Figure 4.7. Grand average source TE matrices pre- and post-TMS to IIPS: Sending and receiving TE between all source pairs after TMS was applied to left intra-parietal sulcus during pre-TMS baseline (top left), 0-200ms (top right), 200-400ms (bottom left), and 400-600ms (bottom right).
Figure 4.8. Grand average source TE matrices pre- and post-TMS to rIPS: Sending and receiving TE between all source pairs after TMS was applied to right intra-parietal sulcus during pre-TMS baseline (top left), 0-200ms (top right), 200-400ms (bottom left), and 400-600ms (bottom right). A relatively large increase in TE, from pre- to post-TMS, can be seen in right parietal sources close to the stimulation site between 0-200ms.
To investigate how total information flow within the brain was affected by TMS, we calculated mean source TE pre- and post TMS to all 4 sites. Mean TE values at each analysis window revealed a consistent pattern: mean TE reached a maximum immediately following TMS and reduced back down towards baseline levels as the trial progressed (see Table 4.1 and Figure 4.9). Paired samples t-tests revealed that TE values within all post-TMS windows were significantly larger than pre-TMS baseline at all 4 sites: p < 0.001 in all tests.

Table 4.1. Mean source TE between windows of interest after TMS to all 4 sites.

<table>
<thead>
<tr>
<th>Stimulation Site</th>
<th>Window</th>
<th>Source TE</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SE</td>
</tr>
<tr>
<td>IATL</td>
<td>Baseline</td>
<td>0.376</td>
<td>0.0054</td>
</tr>
<tr>
<td></td>
<td>0-200ms</td>
<td>0.698</td>
<td>0.0075</td>
</tr>
<tr>
<td></td>
<td>200-400ms</td>
<td>0.531</td>
<td>0.0077</td>
</tr>
<tr>
<td></td>
<td>400-600ms</td>
<td>0.449</td>
<td>0.0056</td>
</tr>
<tr>
<td>rATL</td>
<td>Baseline</td>
<td>0.418</td>
<td>0.0070</td>
</tr>
<tr>
<td></td>
<td>0-200ms</td>
<td>0.668</td>
<td>0.0078</td>
</tr>
<tr>
<td></td>
<td>200-400ms</td>
<td>0.477</td>
<td>0.0072</td>
</tr>
<tr>
<td></td>
<td>400-600ms</td>
<td>0.446</td>
<td>0.0067</td>
</tr>
<tr>
<td>lIPS</td>
<td>Baseline</td>
<td>0.348</td>
<td>0.0061</td>
</tr>
<tr>
<td></td>
<td>0-200ms</td>
<td>0.617</td>
<td>0.0077</td>
</tr>
<tr>
<td></td>
<td>200-400ms</td>
<td>0.396</td>
<td>0.0048</td>
</tr>
<tr>
<td></td>
<td>400-600ms</td>
<td>0.390</td>
<td>0.0062</td>
</tr>
<tr>
<td>rIPS</td>
<td>Baseline</td>
<td>0.372</td>
<td>0.0059</td>
</tr>
<tr>
<td></td>
<td>0-200ms</td>
<td>0.661</td>
<td>0.0077</td>
</tr>
<tr>
<td></td>
<td>200-400ms</td>
<td>0.503</td>
<td>0.0044</td>
</tr>
<tr>
<td></td>
<td>400-600ms</td>
<td>0.404</td>
<td>0.0049</td>
</tr>
</tbody>
</table>
Figure 4.9. Pre- and post-TMS mean source TE within windows of interest: The time-course of information flow around the brain appears to be consistent after TMS was delivered to lATL, rATL, IIPS, and rIPS: TE increases to maximum immediately following the pulse, and reduces back towards baseline in later windows. Error bars indicate ±2 standard error of the mean source TE.

4.4.3 Information flow from TMS-proxy sources

To ascertain which sources were influenced most by the TMS pulse, we analysed the flow of information from the 4 sources that were closest to sites of stimulation to the rest of the brain. In essence, the source magnitude time courses that were nearest to the target sites became proxies for the TMS-evoked activation. The group means of sending TE from the TMS-proxy sources towards the rest of the brain in the -300-100ms window were taken to represent the information outflow from each stimulation site pre-TMS. For example, pre-TMS influence from the IATL proxy source is the 38th row of the top-left matrix in figure 4.5. Distributions of pre-TMS sending TE from TMS-proxies were used to threshold the post-TMS sending TE at the same location (see Table 4.2, and Figure 4.10 for an example of the IATL distribution).
Table 4.2. Descriptive statistics of pre-TMS sending source TE at all 4 TMS-proxies.

<table>
<thead>
<tr>
<th>Stimulation Site</th>
<th>Mean sending TE (bits/sec)</th>
<th>Maximum sending TE (bits/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IATL</td>
<td>0.322 (0.11)</td>
<td>0.6534</td>
</tr>
<tr>
<td>rATL</td>
<td>0.3874 (0.14)</td>
<td>0.6615</td>
</tr>
<tr>
<td>lIPS</td>
<td>0.3311 (0.12)</td>
<td>0.7058</td>
</tr>
<tr>
<td>rIPS</td>
<td>0.4220 (0.16)</td>
<td>1.0244</td>
</tr>
</tbody>
</table>

Figure 4.10. Distribution of baseline sending source TE from IATL TMS-proxy source to rest of the brain: Group average of sending TE from the source closest to IATL stimulation to the 107 sources in the rest of the brain during the pre-TMS -300--100ms analysis window were used to form the distribution. Mean sending TE between the IATL TMS-proxy source and the rest of the brain was 0.322 (0.11) bits/sec. The maximum pre-TMS sending TE (red) which was used to threshold the post-TMS analyses following IATL TMS = 0.6534 bits/sec.
We decided that any sending TE values within the post-TMS analysis windows that were larger than the maximum pre-TMS sending TE at each TMS-proxy were assumed to have been significantly influenced by the TMS pulse at that location: $p > 0.01$ uncorrected = 1/107 sources. The patterns of significant post-TMS sending TE within the 0-200ms, 200-400ms, and 400-600ms windows at all 4 sites were overlaid onto the Colin brain to determine where they were located.

In the 0-200ms window, a large increase of information was evident at many locations following both ATL pulses, but the extent of this influence was confined mainly to the ipsilateral hemisphere of target site (Figure 4.11). After IATL stimulation, largest sending TE values were located within sources closest to the target site in anterior and medial temporal areas. This IATL TMS influence extended posteriorly to include all the sources in the left temporal lobe. There was also a strong increase of information sent to left frontal regions, including orbitofrontal cortex and inferior dorsolateral prefrontal cortex, and a small, but significant, influence of the TMS pulse could be seen in right orbitofrontal cortex too. Furthermore, IATL TMS appeared to evoke activity at left inferior occipital locations including posterior fusiform and lingual gyri, and a small section of right inferior occipital influence in the lingual gyrus also. The left occipital influence extended superiorly into sources that were in primary visual cortices. Finally, there appeared to be a small influence of TMS in left supramarginal gyrus around the temporal-parietal border.
Interestingly, the influence of rATL TMS-proxy revealed a structure that was almost the mirror image of that seen at the lATL site. The largest effects could be seen at all temporal sources, right orbitofrontal and inferior dorsolateral prefrontal cortices, and inferior posterior locations such as fusiform, lingual, and inferior occipital gyri. Moreover, this rATL influence seemed to spread further than the information sent from the lATL TMS-proxy into more superior occipital sources located in right middle occipital gyrus. There were also small, but significant, clusters of influence in left orbitofrontal cortex and left medial temporal lobe.

By the 200-400ms window, the influence of the TMS pulses targeted at bilateral ATLs had receded from the most distant locations found between 0-200ms (Figure 4.12). However, a pronounced influence throughout the ipsilateral temporal lobes was still apparent following TMS. Furthermore, larger sending TE from ATL TMS-proxies was still evident at sources within ipsilateral dorsolateral frontal, orbitofrontal, and posterior inferior occipital regions. Interestingly, the pattern of lATL influence seemed to be slightly more extensive than that found following rATL TMS.

In the final 400-600ms window, the influence of rATL had receded to a small cluster of medial temporal sources near to the target site only. The influence of lATL TMS was also restricted to a smaller subset of left medial temporal sources, but it seemed to extend further than the rATL influence to include some more distant temporal sources and inferior occipital sources as well (Figure 4.13).
Figure 4.11. Information sent from left and right ATL TMS-proxies at 0-200ms post-TMS: lATL influence (orange/red) is located across the temporal lobe, and extends anteriorly to inferior frontal areas, and posteriorly to inferior occipital sources. rATL influence (green/blue) is virtually a mirror image, but does exhibit more extensive spread into superior occipital sources. There is some overlap in orbitofrontal and inferior occipital sources. Locations of TMS-proxies are magenta. TE values thresholded at $p < 0.01$ uncorrected.
Figure 4.12. Information sent from left and right ATL TMS-proxies at 200-400ms post-TMS: Patterns of iATL influence (orange/red) and rATL influence (green/blue) are located throughout ipsilateral temporal, inferior frontal, and inferior occipital cortices, but do not extend as far as the previous window. No ATL influence was found at contralateral sources. Locations of TMS-proxies are magenta. TE values thresholded at p < 0.01 uncorrected.
Figure 4.13. Information sent from left and right ATL TMS-proxies at 400–600ms post-TMS: The pattern of TMS influence following right TMS (green/blue) is limited to a small cluster of medial temporal sources located near to the stimulation site. Left ATL influence (orange/red) has also receded to ipsilateral medial temporal sources, but does extend to incorporate some posterior temporal and inferior occipital sources. Locations of TMS-proxies are magenta. TE values thresholded at $p < 0.01$ uncorrected.
Like ATL stimulation, the patterns of influence following TMS to bilateral IPS in the 0-200ms window were also dominated by activity in the same region that was targeted (Figure 4.14). An increase of information sent from lIPS was apparent in all left parietal sources, and extended into right superior parietal sources as well. Increased proxy-TMS sending TE was also found anterior to the stimulation site in left, and to a lesser extent, right precentral sources. The influence of TMS also spread posteriorly into left, and right, superior occipital gyri, and left middle occipital gyrus. Unlike the temporal stimulation sites, no significant influence of the pulse was found at inferior medial occipital sources, such as the fusiform. Sources that were located laterally to the stimulation site in left superior posterior temporal regions also showed some effects of the TMS pulse. Finally, a small cluster of lIPS TMS influence was found in a source located in left hippocampal gyrus. Once again, the right parietal results mirrored those found after stimulation to the homologue. The influence of rIPS TMS was predominantly found in the same hemisphere and the largest effects were located throughout sources in right parietal lobe. The occipital influence was again visible in bilateral superior occipital gyri, and it extended to the right middle occipital gyrus too. Unlike left parietal TMS, no influence was found in ipsilateral temporal sources. However, the activity that extended anteriorly into frontal regions extended further than that seen following left parietal TMS to include bilateral precentral gyrus and even right superior frontal sources.

By 200-400ms following parietal TMS, the effects of the pulse were only visible in a few sources located near to the respective target sites in both parietal lobes, and inferior occipital cortices (Figure 4.15). The patterns of parietal influence had further reduced in size by 400-600ms such that only superior left occipital sources were visible following lIPS stimulation and the influence of rIPS TMS was reduced to one source in right medial parietal lobe (Figure 4.16).
Figure 4.14. Information sent from left and right IPS TMS-proxies at 0-200ms post-TMS: lIPS influence (orange/red) is located across the left parietal lobe, and extends anteriorly to superior frontal areas, posteriorly to superior occipital sources, and inferiorly to a section of left hippocampal gyrus. lIPS also influences some contralateral sources in right medial parietal, superior occipital, and superior frontal sources. rIPS influence (green/blue) is virtually a mirror image, but does exhibit more extensive spread inferiorly in medial occipital sources. There is some overlap at medial parietal and superior occipital sources. Locations of TMS-proxies are magenta. TE values thresholded at $p < 0.01$ uncorrected.
Figure 4.15. Information sent from left and right IPS TMS-proxies at 200-400ms post-TMS: Left (orange/red) and right (green/blue) TMS influence is localised to sources in medial parietal and medial superior occipital regions within the same hemisphere as the target sites. Locations of TMS-proxies are magenta. TE values thresholded at $p < 0.01$ uncorrected.
Figure 4.16. Information sent from left and right IPS TMS-proxies at 400-600ms post-TMS: The influence of rIPS TMS (orange/red) is now located entirely within a few superior occipital sources. There is one source located in medial parietal lobe still exhibiting lIPS influence (green/blue). Locations of TMS-proxies are magenta. TE values thresholded at $p < 0.01$ uncorrected.
4.5 Discussion

The grand average matrices in the pre-TMS windows at all 4 stimulation conditions had structure, and the largest concentration of information movement seemed to be found within bilateral occipital sources. This is an encouraging first result because our participants would mainly have been processing the visual stimuli projected on screen during this time window. Immediately following stimulation, we observed a large increase of information sent and received throughout the brain. We found a typical looking TEP in our channel data with large deflections relative to baseline (Kähkönen et al., 2005). The component with the largest amplitude peaked at ~150ms, which is within the 0-200ms post-TMS window that gave us our maximum TE values. This gives us some confidence in our analysis. TMS is known to induce larger EEG amplitudes relative to sham stimulation (Zanon et al., 2013), and this was repeated following the increase in TE values immediately post-TMS, total information flow around the brain appeared to reduce down towards pre-TMS levels by the 400-600ms.

Mapping the influence of our TMS-proxy sources on the rest of the brain revealed some promising results. The spread of influence following TMS to all 4 stimulation sites was greatest within the ipsilateral hemisphere, with some contralateral effects being found in plausible areas when you consider the anatomical literature. Moreover, the largest and most long-lasting effects were often found in the same brain regions as the stimulation site. This is an entirely predictable result because local brain areas are highly interconnected by short association fibres that connect adjacent gyri (Catani et al., 2012).

Both left and right ATL TMS seemed to produce a broadly similar pattern of influence that mirrored each other. The temporal lobes are highly connected regions of the brain that contain six major association fibres and two main commissural fibres (Bajada et al., 2016). A substantial increase of information from ATL TMS-proxies could be seen in the ipsilateral orbitofrontal cortex, which could have spread via the uncinate fasciculus that
connects these two locations (Catani et al., 2002). The ATL evoked increase of TE found in ipsilateral inferior dorsolateral prefrontal cortex could be as a result of short association fibres connecting them to orbitofrontal areas, or by the arcuate fasciculus which is part of the superior longitudinal fasciculus and runs from posterior superior temporal regions to dorsal prefrontal cortex around Broca’s area (Catani et al., 2002; Makris et al., 2005). The ipsilateral inferior occipital influence could have spread via the inferior longitudinal fasciculus (Catani et al., 2003). Both ATL TMS-proxies were found to influence contralateral orbitofrontal sources, and this result could have been achieved via the anterior commissure which connects the two hemispheres and has projections into both areas (Catani et al., 2002). Therefore, the anterior commissure could also be responsible for the increased TE in the left medial temporal lobe that was evident after right ATL TMS. Finally, the left ATL TMS seemed to influence a section of supramarginal gyrus in the parietal lobe, which could be accounted for by the superior longitudinal fasciculus (Makris et al., 2005). Furthermore, these locations converge neatly with semantic task-related brain activations found in studies that utilised both PET (Devlin et al., 2000) and fMRI (Jung and Lambon Ralph, 2016) to study neural activity in the ATLs. It is logical to assume that a network of brain regions that activates during a task will be highly connected. Furthermore, neural activity evoked at a node within this network will influence activity throughout its connections. The similarities between our map of IATL TMS influence and the patterns of neuroimaging activation during tasks that probe the function of IATL gives us some confidence that our analysis is really tracking the TMS pulse.

Like our temporal results, the patterns of influence identified after left and right parietal TMS were broadly mirror images of each other. The maps created following parietal TMS located the influence primarily in bilateral parietal, superior occipital, and superior prefrontal sources. Furthermore, the extent of influence seemed to be slightly greater in the hemisphere ipsilateral to stimulation site. This pattern of activity can be accounted for by the extensive network of association and callosal fibres that exist between these
regions (Catani et al., 2002; Wakana et al., 2004). Additionally, prefrontal and parietal cortices are connected via the SLF1 component of the superior longitudinal fasciculus (Makris et al., 2005). Interestingly, the left IPS stimulation produced some influence in left temporal sources located in parahippocampal gyrus. Theoretically, this could have been achieved via the posterior branch of the cingulum which connects the parietal lobe to the parahippocampal gyrus. We didn’t find the same medial temporal influence following right parietal stimulation, but within the map following rIPS TMS we can see a shape that resembles the cingulum’s characteristic hook around the genu and splenium as it projects down towards the parahippocampal gyrus (see bottom left brain slice in Figure 4.16: Catani and Thiebaut de Schotten, 2008).

4.5.1 Conclusion
In conclusion, we have presented here an analysis method that appears to track the causal interactions that arise when a brain region is stimulated by TMS. Using our paradigm, stimulation at 4 different sites produced the same effects across the entire brain: the amount of information being exchanged reached maximum during the first 200ms post-TMS, and decreased below the levels found at pre-TMS baseline between 400-600ms after a pulse. When we compare the areas in which we find the largest influence of the TMS pulses we find that they correspond favourably with the anatomic tractography literature; we even find that we can trace this influence through shapes that resemble known tract structures. With some more refinement, this method might become a useful addition to the battery of techniques used to assess cortical connectivity. We have only scratched the surface of the technique’s potential.
Chapter 5 – Discussion

There were two main aims of this thesis: the first was to create a novel method of mapping effective connectivity (EC) in the human brain by analysing electroencephalography (EEG) source data with trial-based transfer entropy (TE); and the second was to use the technique to investigate brain function during semantic and number processing.

Chapter 1 introduced the concepts of functional and effective connectivity, and discussed how the research field of cognitive neuroscience has used neuroimaging techniques to investigate them. Electrophysiological imaging methods, such as electroencephalography (EEG) and magnetoencephalography (MEG), are direct measures of brain activity and as such are optimal for investigating inter-cortical connectivity (Gross et al., 2001). Many methods are employed to quantify connectivity within M/EEG data, but not all are sensitive to the direction of causal influence between regions, and some that are rely on the prior definition of models so might not represent true network behaviour. Furthermore, connectivity is often investigated at the sensor-level which reveals nothing about the interactions between cortical regions (Schoffelen and Gross, 2009). Therefore, a better strategy would be to apply connectivity measures to M/EEG source data. An optimal quantifier of connectivity will be: data-driven to enable exploratory analyses; sensitive to both linear and non-linear interactions; and sensitive to the direction of these interactions over a range of time delays (Vicente et al., 2011). The information theoretic metric Transfer entropy (TE) satisfies all these criteria. Theoretically, analysing time-locked M/EEG source data with TE could elucidate the real-time dynamic connectivity between brain areas non-invasively, and would be a useful addition to the battery of methods employed to study connectivity. This thesis documents the creation and application of just such a method.
This chapter is arranged into 4 sections: the first will summarise the findings of the methods and experimental chapters presented in this thesis; the second section will discuss some outstanding methodological issues that require investigating; the third section will suggest further developments to the method that could improve it; and in the final section there will be a discussion of the broader implications of the analysis method to brain function research.

### 5.1 Review of findings

In Chapter 2, EEG data recorded during a semantic category judgement task and a number magnitude judgement task were analysed using traditional event-related potential (ERP) techniques. As hypothesised, the well-known N400 component was visible in the ERP after semantic judgments were made (Kutas and Hillyard, 1980a), and a slightly earlier component (termed the N365) was found to contribute to number magnitude processing. Each trial consisted of 4 sequential stimuli presentations: 3 priming and 1 target. N400 amplitudes were found after all 4 words appeared on screen, and the amplitudes reduced sequentially with the first being the largest and the last being the smallest. Similar reductions to repeated neural activity such as this are common to priming paradigms (Henson and Rugg, 2003). Brain activity in a time-window which incorporated the peak of the post-target N400 was localised to left anterior temporal lobe (ATL) and left inferior frontal areas which compared well with the locations found in the combined PET/fMRI study from which the task was adapted (Devlin et al., 2000). Much evidence exists to suggest that the ATLs act as an area where sensory information from all modalities converges to form semantic concepts (Patterson et al., 2007; Lambon Ralph, 2014). Brain activity associated with the N365 component was predominantly localised to right superior parietal regions including the intra-parietal sulcus, which has been proposed as the location of a “quantity system” that is involved in mental arithmetic and number comparison (Nieder and Dehaene, 2009).
In Chapter 3, the same data used in the previous chapter was analysed at the source level with the trial-based TE method which is a focus of this thesis. A moving analysis window was passed over the data, and patterns of information flow at 0-200ms, 200-400ms, and 400-600ms following both word and number targets were identified. In the earliest windows, the patterns of information exchange were concentrated between bilateral occipital sources following both targets. By the latest window, more wide-spread interactions across the whole brain were evident. The presentation of visual stimuli in EEG experiments reveals a common set of components called the visual-evoked potential (Jeffreys and Axford, 1972), which is thought to be generated by activity in occipital sources (Di Russo et al., 2002). Semantic processing is believed to start with feature extraction of stimuli in sensory regions of the brain, before progressing to wide-spread areas across the brain including the ATLs between ~200-400ms (Halgren et al., 2002; Marinkovic et al., 2003). This was echoed by the TE analysis which found the largest amounts of information exchange occurred within the 200-400ms time-window (see Figure 3.6). The TE analysis was able to detect significant differences of information exchange in both word and number data between particular brain regions. At 0-200ms there was significantly more information moving posteriorly from temporal sources to occipital regions for words than numbers. This result could provide support for a top-down facilitation model of semantic processing, which suggests that some early semantic processing from frontal and temporal areas can be back-projected to posterior sensory regions as an “initial guess” to aid bottom-up processes (Bar, 2006). Significantly larger TE values were found after words in the direction from occipital sources to parietal regions between 0-200ms. Parietal activity related to the processing of semantic word properties was found around 200ms in an EEG investigation of visual word recognition (Hauk et al., 2006). Increased influence from frontal to temporal sources was found for numbers relative to words between 0-200ms, and could be representative of the extra WM demands required for word task performance (Gevins et al., 1997). By 200-400ms the information exchange between frontal and temporal sources was significantly larger in both directions following number targets. The medial temporal lobe is known to be
active during WM tasks, and WM load has been found to increase proportionally with MTL gamma activity (Ranganath and D'Esposito, 2001; Nichols et al., 2006; Axmacher et al., 2007). The presentation of number targets also resulted in significantly more information moving from frontal sources towards parietal regions between 200-400ms. Furthermore, information exchange between bilateral parietal sources was significantly larger following numbers than words at 200-400ms too: a result which complements the localisation of the N365 number magnitude judgement component in Chapter 2. Taken together, this could be evidence of the timings within a distributed network of frontal, temporal, and parietal sources combining to complete the magnitude judgment: early frontal expectations about the number target could trigger retrieval of the 3rd number at medial temporal sources between 0-200ms, this information is conveyed back to frontal sources before combining with knowledge of the sequence rule at parietal sources where the judgment is made at ~365ms.

Neuroimaging techniques often investigate brain function by manipulating some facet of an experimental task, and use statistical inference to identify the network of connected brain areas responsible for behavioural performance. A less ambiguous method of tracking effective connectivity would be to perturb the brain, using transcranial magnetic stimulation (TMS), at a region and try to identify effects of the evoked neural activity across the brain. In Chapter 4, the TE analysis was applied to combined TMS-EEG data to track the influence of the stimulation as it spread from 4 different target sites to inter-connected brain regions. Relative to pre-TMS baseline levels, information flow across the brain was greatest in the 200ms analysis window immediately following TMS at all 4 sites, and by 600-800ms it had reduced back down towards baseline. This result could have been predicted, as the largest deflections in the ERP were found in a consistent sequence of components during the first 200ms after stimulation that resembled characteristic TMS-evoked potential (Kähkönen et al., 2005; Miniussi and Thut, 2010).

The influence of the TMS pulse throughout the brain was assessed by analysing the information sent from the source closest to each site of TMS stimulation. Compared to
pre-TMS baseline, the largest and longest lasting effects were found at sources located in the same region as the TMS target, and in general the influence of stimulation was mainly found at regions in the same hemisphere. Following TMS to left anterior temporal lobe (lATL), a significant rise of TE was found in left temporal, inferior frontal, inferior occipital, and inferior parietal locations. Furthermore, there was some influence found in the right hemisphere within orbitofrontal, and lingual gyrus. Right ATL stimulation resulted in a mirror image of lATL influence: ipsilateral wide-spread temporal, inferior frontal, and inferior occipital sources. However, the contralateral influence was slightly different as some was found in a left medial temporal source, and there was more right inferior occipital influence relative to lATL TMS. All the sources where temporal TMS influence was found could be predicted by the tract configurations of the temporal lobe (Bajada et al., 2016). Temporal TMS influence could spread via short association fibres, inferior frontal via uncinate fasciculus (Catani et al., 2002), inferior occipital sources via inferior longitudinal fasciculus (Catani et al., 2003), inferior parietal sources via the superior longitudinal fasciculus (Makris et al., 2005), and contralateral orbitofrontal via the anterior commissure (Catani et al., 2002). Intra-parietal sulcus (IPS) TMS resulted in broadly similar patterns of influence in both hemispheres: widespread bilateral parietal, bilateral prefrontal, bilateral premotor, and widespread bilateral occipital sources in superficial cortex. The influence at prefrontal, premotor, and occipital regions showed greatest extent ipsilateral to stimulation site. Once again, these patterns of influence could be explained quite easily by the known tract anatomy. Short association fibres can account for the ipsilateral and contralateral parietal influence, and the spread posteriorly into superior occipital sources and beyond. The TMS influence could have reached prefrontal and premotor regions via the SLF1 component of the superior longitudinal fasciculus (Makris et al., 2005). Interestingly, following left IPS there appeared to be some influence at left hippocampal gyrus, which could have reached there by the ventral projection of the cingulum (Catani and Thiebaut de Schotten, 2008).
5.2 Outstanding methodological issues

The analysis as presented has had some success at determining the causal interactions that exist within semantic and number processing networks, or the influence of a TMS pulse as it spread to inter-connected regions from the stimulation site. However, there are some outstanding methodological issues that require investigating to improve the method, and they are discussed in this section. The topics covered in this section are: (1) the effects of window size; (2) alternative source variables to analyse; (3) source grid resolution; and (4) identifying sources of interest in group level data.

5.2.1 The effect of window size

The analysis windows employed in both instances of the TE analysis have been 200ms in length. More work needs to be done on the effects of shortening the window length, because this will allow us to better visualise the dynamic interactions within brain networks. However, there will be a limit on the minimum window size with which we can still obtain unbiased TE estimations. Any biasing of shortened window TE estimations will likely be caused by the concept of information rate. Information rate is the average entropy per symbol in a stochastic process, and it has been shown to increase as the time resolution of an analysis decreases (Strong et al., 1998). Informally, information rate can be thought of as the time density of the average information of a process. As our analysis windows get smaller, the variability of values in our windows and between our windows will become smaller, and therefore the distributions of outcomes will become more uniform. In effect, the random variables in our system will become more dependent and begin to form a stationary process. There will be an optimum window length using our analysis in its current form, but it will only be found with more investigation.

5.2.2 Alternative source variables

The time series of source magnitude estimations obtained from the LCMV beamformer were 3-dimensional moments at each time point. In the TE analysis, the magnitude was
projected in its strongest direction using the Euclidean norm. However, it is highly unlikely that the direction of each source remained static over the course of the trial length using this method. It should be possible to account for the variability of source direction in this analysis by incorporating it into our calculation of TE too. Furthermore, it might be possible that a better metric to run the analysis on is the neural activity index (NAI) rather than the source magnitude itself. NAI is an estimate of the source to noise variance as a function of location (Van Veen et al., 1997). NAI calculation requires knowledge of the noise covariance matrix, and essentially uses this estimate of spatially inhomogeneous noise to normalise the source power. NAI is supposed to represent the true source positions better, and is often used to correct for a centre of the head bias that results from a noise bias in source reconstructions produced by beamformers.

5.2.3 Source grid resolution
The source grid used in the analysis had a resolution of 20mm between points that is too coarse to make precise claims about the structure of the networks influenced by TMS or our task manipulations. The grid was created to try and achieve the maximum cortical coverage, using as few sources as possible equally located in both hemispheres. These criteria were imposed because trial-based TE estimations were found to be time consuming, so there was always a trade-off between practical estimation duration and grid resolution. This problem was further exacerbated during the surrogate significance testing methodology because we could be running the analysis on 1000 datasets per participant to build up a large enough distribution of surrogate values. As processing power increases, it should be possible to create higher resolution grids and still run the analysis

5.2.4 Identifying sources of interest in group level data
After creating grand average TE matrices, we determined whether there were significant differences between experimental conditions, or successive time-windows in the same condition, by averaging TE values at each source, or within brain regions, before testing for significance using appropriate statistical techniques. This technique is sound, but the
results only provide us with a superficial knowledge of information flow across or between regions, or the entire brain. For example, the connectivity profile of the temporal lobe was assessed as a whole, even though we know that posterior, anterior, and medial temporal areas are responsible for very different sets of functions. Ideally, the method should be able to identify individual sources, or clusters of sources (if the resolution is improved), that possess similar connectivity profiles, and, by extension, start to map the networks that are preferentially activated by an experimental task. Once these sources of interest have been identified from the data, then you could perform more detailed analyses as described in the further developments section.

**5.3 Further developments**

Once the methodological issues are resolved, the trial-based TE analysis could be augmented with some developments that could make the technique more powerful. The following section details two theoretical expansions to the method, and an idea for a new analysis that was influenced by the results of the TMS-EEG experiment. These are: (1) a source of interest analysis; (2) combing the TE analysis with tractography data; and (3) TE analyses of TMS paradigm manipulations.

**5.3.1 Sources of Interest Analysis**

Once the sources of interest (SOI) have been determined in the matrices, an alternative TE analysis could be performed to ascertain whether these connections exhibit a particular timing structure. TE could be estimated across windows that overlap by a small amount, and rather than do an average over a range of time delays, TE could be plotted as function of time delay across the trial length. Intuitively, the length and direction of a connection between sources will affect the time taken for activity in a sending source to reach its destination. By calculating how the magnitude of TE varies with time delay length it should be possible to determine the optimal time delay with which to study connectivity between SOIs.
5.3.2 Refining connectivity maps with tractography data

The connectivity maps produced by tracking a TMS pulse through the brain appeared to be well explained by the anatomic connectivity literature. To further refine the TE maps it should be possible to incorporate an individual’s anatomical connectivity profile as described by tractography data. Tractography is a method that uses diffusion tensor imaging (DTI) to reconstruct the white-matter architecture via the direction of water diffusion through the brain (Basser et al., 2000). White-matter tracts are required to transmit neural activity between brain regions, so there is an obvious link between effective and anatomical connectivity. A simple method of combining these two types of data would be to limit the TE analysis to source interactions that tractography deems are likely connected. A slightly more involved approach would be to weight the TE values as a function of the tractography value.

5.3.3 TE analysis of TMS paradigms

A key result of the TMS-EEG experiment was the change of gross information flow across the entire brain as the trial progressed. The largest TE values were found immediately following the pulse which would suggest that the pulse influenced activity at many locations throughout the brain. TMS-EEG investigations of channel-level connectivity have found that the amplitude of the TEP increases linearly with stimulation intensity (Komssi et al., 2004), and it would be interesting to see whether it had a similar effect on the magnitude of TE. Furthermore, our TE analysis could potentially be used to investigate cortical excitability following popular TMS paradigm manipulations such as intensity and coil position, because the effects of TMS on the brain are still not fully understood (Siebner et al., 2009).
5.4 Broader implications

Brain connectivity can be investigated via multiple imaging modalities and analysed using many techniques, but all these methods have their strengths and weaknesses. The novel method produced in this thesis combined EEG, TMS, and TE to map the transient activity within brain networks. It was inspired by the current thinking within cognitive neuroscience concerning the optimal strategies for investigating connectivity: we used a direct measure of neural activity that is able to capture brain function with excellent temporal resolution (EEG); we perturbed the brain and looked for effects of the evoked neural activity (TMS); and we used a quantifier of effective connectivity that is suitable for the exploratory analyses of linear and non-linear influence over a range of time-scales (TE).

The key strength of the technique is its ability to capture the timing and direction of neural activity as it moves through the brain networks. Some methods can determine the when brain regions are functionally connected at millisecond time-scales. For example, we could have used dynamic imaging of coherent sources (DICS: Gross et al., 2001) to determine that frontal sources were functionally connected to temporal sources during the first 200ms in our magnitude judgment task. However, using our technique we could show that frontal sources were influencing temporal sources in this time window, before a two-way exchange of information manifested between the regions at 200-400ms. Furthermore, we found some evidence that supports early top-down influence of semantic activity at temporal sources influencing posterior sensory regions between 0-200ms. Once the identified methodological issues have been thoroughly investigated, the method could potentially make an important contribution to our models of semantic processing.
The potential for investigations of cognitive function using this method are excellent. For instance, it could theoretically be used to investigate the disordered connectivity between regions that characterise disorders such as schizophrenia (Skudlarski et al., 2010). Furthermore, the method appears able to do this in vivo using equipment that is readily available to universities and hospitals, and at a fraction of the cost of some other neuroimaging techniques.

5.5 Conclusion

After many years of studying brain function, we are only beginning to understand how brain regions work together to perform cognitive processes and influence human behaviours. This thesis created a combined technique for investigating brain connectivity, which showed promise in being able to map transient interactions between interconnected areas of the brain. Using the technique, we have been able to track the spread of evoked-neural activity as it propagates from the site of stimulation to the brain regions with which it is functionally connected. We were also able to find task-dependent differences in the direction of information flow between brain regions whilst participants performed semantic and number processing. One day, science might finally determine how brain function arises from the neural activity we can visualise using neuroimaging techniques, and methods such as this might make that time come sooner.


associated functions.’ Cortex.


Kutas, M. and Federmeier, K. D. (2011) 'Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP).'</
psychology}, 62, January, pp. 621–47.


studies of cognition.’ Cortex; a journal devoted to the study of the nervous system and behavior, 45(9) pp. 1035–42.


