# Bayesian network modeling of neural systems with functional MR images

Zheng, Xuebin

2007

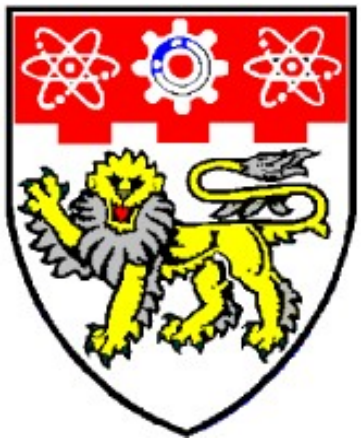
Zheng, X. B. (2007). Bayesian network modeling of neural systems with functional MR images. Doctoral thesis, Nanyang Technological University, Singapore.

https://hdl.handle.net/10356/2585

https://doi.org/10.32657/10356/2585

Nanyang Technological University

# BAYESIAN NETWORK MODELING OF NEURAL SYSTEMS WITH FUNCTIONAL MR IMAGES



**ZHENG XUEBIN**

**SCHOOL OF COMPUTER ENGINEERING**

**NANYANG TECHNOLOGICAL UNIVERSITY**

**2007**

# Bayesian Network Modeling of Neural Systems with Functional MR Images

**Zheng Xuebin**

**School of Computer Engineering**

A thesis submitted to the Nanyang Technological University
in fulfilment of the requirement for the degree of
Doctor of Philosophy

**2007**

# Table of Contents

# List of Tables

# List of Figures

# Acknowledgements

I would like to express my sincere gratitude to Associate Professor Jagath C. Rajapakse, my supervisor, for his many suggestions and invaluable guidance throughout my research. I have learned a lots of things from him, ranging from academic knowledge to research experiences. Without his help, my research would not have reached this stage.

All the members of BioInformatics Research Center should be thanked for their friendship and creating a very pleasant research environment.

Besides, I would enthusiastically thank my family and my beloved Yannie, for their constant support for my work and warm care for my life.

Finally, many thanks to School of Computer Engineering, Nanyang Technological University for offering me the award of Research Scholarship and providing first-rate research facilities.

# Abstract

Neuroscientists have shown an increased interest in knowing interactions among brain regions activated during sensory and cognitive tasks in order to obtain a holistic view of brain function. The existing methods of connectivity analysis, such as Structural Equation Modeling (SEM), Dynamic Causal Modeling (DCM), and Granger Causality Mapping (GCM), are confirmatory in the sense that they need a prior connectivity model to begin with. These methods are often under anatomical constraints or complicated by the fact that many of them have been obtained in the studies of monkeys. It is not always certain which areas are to be included in the studies, especially if the brain regions involved are specific functions that are unique to humans such as language and cognition.

This thesis presents an exploratory (data-driven) approach based on Bayesian networks in modeling neural systems with functional MR images. Bayesian networks are directed graphs where the effective connectivity between two brain regions is represented by conditional probability densities (CPD). Therefore, the interactions in the network are represented in a complete statistical sense. The results with synthetic data show that Bayesian networks can better fit the imaging data obtained in functional MR experiments. The effective connectivity patterns obtained from silent reading tasks are consistent with previous literature. The method was further applied on fMRI data collected in English letter searching tasks for deriving a plausible model of language system.

Disconnectivity among brain regions is hypothesized in brain diseases such as dyslexia, Alzheimer's disease, Parkinson's disease, and schizophrenia. The previous

methods of testing disconnectivity hypotheses were mostly done by comparing the activation patterns. However, this is not effective when the diseases are due to deficits in connectivity. This thesis demonstrates how the graphical models derived by the present method can be effectively used in the analysis of lesion studies, using a counting Stroop task and a case study of a stroke patient. These studies indicate the promise of the present method as a general framework for analyzing a wide range of brain disorders in future.

# Author's Publications

## Journal Papers

1. Jagath C. Rajapakse, Choong Leong Tan, Xuebin Zheng, Susanta Mukhopadhyay, Kanyan Yang. Exploratory analysis of brain connectivity with ICA. *IEEE Engineering in Medicine and Biology Magazine*, March/April Issue, Vol. 25, No. 2, 2006, pp. 102-111.

2. Xuebin Zheng and Jagath C. Rajapakse. Learning functional strucutre from fMRI images. *Neuroimage*, Vol. 31, No. 4, pp. 1601 - 1613, July 2006.

3. Jagath C. Rajapakse, Xuebin Zheng, Susan J. Rickard Liow, Keren-Happuch E, Winston E.H. Lim. Letter search in high and low frequency words: neural systems study with fMRI. *Human Brain Mapping*, (under review).

4. Jagath C. Rajapakse, Xuebin Zheng, Susan J. Rickard Liow, Keren-Happuch E, Winston E.H. Lim. Letter search in legal and illegal nonwords: neural systems study with fMRI. *Human Brain Mapping*, (under review).

5. Jagath C. Rajapakse, Yang Wang, and Xuebin Zheng. Probabilistic framework for effective brain connectivity from functional MR images. *IEEE Transactions on Medical Imaging*, (under review).

## Other Journal Publications

1. Xuebin Zheng, Winston E.H. Lim, Susan J. Rickard Liow, and Jagath C. Rajapakse. Modeling functional connectivity of brain for lesion study. *Neuroimage*, Volume 31, Sup. 1, S62, June 2006 (abstract).

2. Jagath C. Rajapakse, Susan J. Rickard Liow, Keren-Happuch E, Winston E.H. Lim, Xuebin Zheng, Lynn G. Ho, Wendy W.P. Tham, Choong Leong Tan. Strong vs weak lexicalisation: letter search in high and low frequency words. *Neuroimage*, Volume 31, Sup. 1, S39, June 2006 (abstract).

3. Jagath C. Rajapakse, Susan J. Rickard Liow, Keren-Happuch E, Winston E.H. Lim, Xuebin Zheng, Lynn G. Ho, Wendy W.P. Tham, Choong Leong Tan. Orthographic processing: letter search in legal and illegal nonwords. *Neuroimage*, Volume 31, Sup. 1, S56, June 2006 (abstract).

## Conference Papers

1. Xuebin Zheng and Jagath C. Rajapakse. Graphical models for brain connectivity from functional imaging data. *International Joint Conference on Neural Network* (*IJCNN*), Hungary, 2004.

2. Xuebin Zheng, Jagath C. Rajapakse. Disconnectivity analysis of brain disease from fMR images. *International Conference on BioMedical Engineering* (*ICBME*), Singapore, August 2005.

# Abbreviations and Symbols

## Abbreviations

| | |
|---|---|
| 3-D | three-dimensional |
| ACC | anterior cingulate cortex |
| AFG | anterior frontal gyrus |
| ASL | American sign language |
| BA | Broadmann areas |
| BIC | Bayesian information criterion |
| BNs | Bayesian networks |
| BOLD | blood oxygen level dependent |
| CPD | conditional probability densities |
| CSF | cerebrospinal fluid |
| DAG | directed acyclic graph |
| DCM | dynamic causal modeling |
| DTI | diffusion tensor imaging |
| EC | extrastriate cortex |
| EPI | echo planar imaging |
| FHWM | full width at half maximum |
| fMRI | functional magnetic resonance imaging |
| FUG | fusiform gyrus |
| FWER | family-wise-error-rate |
| GCM | Granger causality mapping |
| GLM | general linear model |
| GRF | Gaussian random field |
| HFW | high frequency words |
| HRF | hemodynamic response function |

| | |
|---|---|
| IC | interference-control |
| ICA | independent component analysis |
| ICBM | international consortium for brain mapping |
| IFG | inferior frontal gyrus |
| IFX | interference-fragile X |
| INW | illegal non-words |
| IPL | inferior parietal lobe |
| LFP | local field potential |
| LFW | low frequency words |
| LNW | legal non-words |
| LR | likelihood-ratio |
| MAP | maximum a posteriori |
| MCMC | Markov chain Monte Carlo |
| MFG | middle frontal gyrus |
| MH | Metropolis-Hastings |
| ML | maximum likelihood |
| MNI | Montreal neurological institute |
| MRFs | Markov random fields |
| MTC | middle temporal cortex |
| NC | neutral-control |
| NFX | neutral-fragile X |
| PCA | principal component analysis |
| PMA | primary motor area |
| rCBF | regional cerebral blood flow |
| SEM | structural equation modeling |
| SMA | supplementary motor area |
| SPL | superior parietal lobe |
| SPM | statistical parametric mapping |
| VAR | vector autoregressive |
| VIFG | ventral inferior frontal gyrus |
| VST | visual search task |

## Important Symbols

| | |
|---|---|
| $b_{ij}$ | linear coefficient between region $i$ and $j$ |
| $c_{ij}$ | pairwise correlations between region $i$ and $j$ |
| $C$ | correlation matrix |
| $D$ | dataset |
| $e$ | Gaussian innovation |
| $G$ | design matrix in a GLM |
| $M$ | linear coefficient parameter matrix |
| $N$ | size of a dataset |
| $p(a\|b)$ | conditional probabilities of $a$ given $b$ |
| $p(x)$ | joint probability of the activations |
| $r_i$ | $i$th brain region |
| $R$ | a set of brain regions |
| $S$ | network structure |
| $W$ | precision matrix of a normal distribution |
| $x_i$ | time-series of the $i$th activated brain region |
| $\beta$ | regression parameter in a GLM |
| $\theta$ | parameters of the CPD |

# Chapter 1

# Introduction

Functional Magnetic Resonance Imaging (fMRI) is increasingly utilized in the exploration of brain networks and neural interactions underlying brain functions. A variety of analysis methods have been developed for detecting brain activations with fMRI. The analysis of an fMRI image series is most often based on the computation of Statistic Parametric Maps (SPM) and making inferences from that. SPM is computed by adapting a general linear model [4]. Non-parametric approaches such as Kolmogorov-Smirnov [5] statistics have also been used. Inferences from SPMs are usually based on voxel-wise intensities above a threshold [6], on the spatial extent of contiguous voxels above a threshold (cluster-size thresholds) [7, 8] or on the combination of these two [9]. The parameters of tests are chosen so that the probability of detecting a false activation in the whole volume (or in the search volume) is relatively small, e.g. 0.05. Gaussian random field (GRF) [10], Markov random field (MRF) [11], and conditional random field (CRF) [12] have been used for correction of contextual information of data. Model free data-driven methods such as principal component analysis (PCA) [13], independent component analysis (ICA) [14], and fuzzy $c-means$ clustering [15] have also been used for analysis of fMRI data [16].

Although many researchers have attempted to identify individual brain areas involved in various cognitive tasks, holistic views of effective connectivity of higher-order functions have not been investigated thoroughly. More recently, functional integration studies describing how functionally specialized areas interact and how these interactions lead the brain to perform a specific task have become one of the hot topics in brain mapping research [17].

Presently, information about neural interactions is often extracted by decomposing interregional covariances among activations. Structural Equation Modeling (SEM) has been the most commonly method used to analyze the effective connectivity among brain regions. The covariances among the brain regions in SEM describe the behavior of a neural system in a second-order statistical sense whereas the conditional probability densities (CPD) characterising Bayesian network describes the behavior of a network in a complete statistical sense. Dynamic Causal Modelling (DCM) was introduced by Friston [3] to model functional interactions at the neuronal level and comprises of a bilinear model for neurodynamics and an extended balloon model for hemodynamics. DCM models interactions at the neuronal rather than the hemodynamic level [17], which is more useful in analyzing the temporal interactions between brain regions. Instead, the present approach focuses on exploring the structure of interactions of the neural systems. Granger causality mapping (GCM), a linear method developed for modelling time-resolved fMR time-series, investigates effective connectivity among activated brain areas by using a vector autoregressive (VAR) model [2]. Granger causality mapping renders a voxel-wise connectivity analysis whereas the present approach seeks for a global representation of a neural system with a region-wise connectivity.

The present method based on Bayesian networks allows extraction of the connectivity among brain regions from functional MRI data in an exploratory manner. Bayesian network modelling is widely applicable for compactly representing the joint probability distribution over a set of random variables [18]. In the functional brain networks, the nodes represent the activated brain regions and a connection between two regions represents an interaction between them. The Maximum A Posteriori (MAP) estimation of the structure of the functional network is derived from fMRI data to maximize the Bayesian Information Criterion (BIC) by using a greedy search algorithm.

## 1.1 Motivation

Graphical models, also known as *belief networks* or *Bayesian networks*, are widely applicable formalism for compactly representing the joint probability distribution over a set of random variables. They are a marriage between probability theory and graph theory [18]. They have the following four main advantages [19]:

1. They can readily handle incomplete data sets: When one of the input variables is not observed, most models will produce an inaccurate prediction. Bayesian networks offer a natural way to encode the correlation between the input variables.

2. They allow one to learn about causal relationships: It is useful during exploratory data analysis and allows us to make predictions in the presence of interventions.

3. They facilitate the combination of domain knowledge and data: Prior or domain

knowledge is important especially when data is scarce or expensive. Prior knowledge and data can be combined with well-studied techniques from Bayesian statistics.

4. They offer an efficient and principled approach for avoiding the over fitting of data: There is no need to hold some of the training data for testing.

In this thesis, a novel method is proposed for exploratory way of learning the structure of effective connectivity among brain regions involved in a functional MR experiment, using Bayesian networks. The approach is exploratory in the sense that it does not require an *a priori* model as in the earlier approaches. In a Bayesian network, the activated brain regions are represented by the nodes in a directed acyclic graph (DAG) and the interactions among them are represented by connections among the nodes. The Maximum A Posteriori (MAP) estimation of the structure of the functional network can be derived from fMRI data to maximize the Bayesian Information Criterion (BIC) of a DAG. The Metropolis-Hastings (MH) algorithm [20], a Markov Chain Monte Carlo (MCMC) method, can be used to search the space of DAGs to find the optimal structure of the network.

The last decade has witnessed many new and non-invasive neuroimaging techniques (e.g., fMRI, PET, EEG and MEG) allowing researchers to investigate the specific brain regions involved in higher-order brain functions such as language processing. However, elaborate neuroimaging studies that describe how brain regions interact in exploratory way do not exist. A neuroanatomical network model will be proposed for language processing across three main components: orthography, phonology and semantics, by using the present data-driven method. As it may be difficult to truly isolate any particular component in language processing, it is worthwhile to

discuss the model in three components collectively. The method will be applied on the fMRI datasets from language tasks to introduce and discuss a plausible language system model.

The previous approaches mostly detect brain disorders by comparing activation patterns between patients and matched controls. However, such approaches do not reflect the abnormalities due to disconnectivity, such as the functional disconnection of the left angular gyrus involved in normal reading task, that has been found to be implicated in the disease of dyslexia [21, 22], functional disconnectivity of the medial temporal lobe in Asperger's syndrome [23], and the functional disconnectivity in subjects at high genetic risk of schizophrenia [24], etc. The present method of connectivity can be used in the analysis of disorders by studying the differences between the functional networks of patients and healthy participants, derived from fMRI data.

The strength of interactions between regions is another important factor for connectivity analysis. A brain disorder could be due to either a complete disconnectivity or a weak connection. The method could be extended by estimating the parameters of the connections, where the values of conditional probabilities reveal how much an activation of a region depends on the others. Hence, the brain disorders could be explored more completely.

In short, the research has culminated towards the followings:

1. Introduction of Bayesian networks for the effective connectivity analysis of functional MR images.

2. Study of neural systems involved in particular task such as reading, working memory, visual, and motor tasks, etc.

3. Simulate lesion studies, and interpret abnormal behavior.

4. Study disconnectivity hypotheses made in diseases such as dyslexia, Alzheimer's disease, Parkinson's disease and schizophrenia, etc.

## 1.2 Major Contributions

The major achievements of this research are listed as entries below:

- A novel method was proposed by using Bayesian networks to learn the structure of effective connectivity among brain regions involved in a functional MR experiment [25]. The approach is exploratory in the sense that it does not require an *a priori* model as in the earlier approaches, such as SEM or DCM, which can only affirm or refute the connectivity of a previously known anatomical model or a hypothesized model. The conditional probabilities that render the interactions among brain regions in Bayesian networks represent the connectivities in the complete statistical sense. The present method is applicable even when the number of regions involved in the cognitive network is large or unknown. The method was tested and compared with earlier methods by using synthetic data. The results of data simulation show that the method is robust to noise, and performs better than the SEM method as the number of region grows. The performance of the present approach was further demonstrated on fMRI data collected in silent word reading task. The network derived from silent reading data was consistent with the literature. The network demonstrated the dominance of language processing in the left hemisphere and the regions in the right hemisphere receives the effects of processing from the left hemisphere.

- FMRI data was collected in a letter searching task which was designed to compare reading of high frequency words (HFW), low frequency words (LFW), legal non-words (LNW) and illegal non-words (INW). The results reveal the differences between the three components of language processing and yield a plausible model explaining how these regions are connected at the three subcomponents of language processing (orthography, phonology, and semantics). The complexity of the brain makes it difficult to be explored, especially in higher cognitive tasks; the analysis of functional integration (functional connectivity and effective connectivity) is still far from settled. The proposed method of exploring global neural systems from functional imaging data provides an alternate method to study brain function in terms of networks.

- Networks derived from patients and healthy participants were used to explore disconnectivities of brain diseases. The basic idea is that networks learnt from a patients and a normal person doing the same tasks show how the connectivity differs from one another. Hence functional disorders could be diagnosed. This approach is a global disconnectivity analysis, which benefitted from the characteristic of the complete statistical analysis. The approach was demonstrated by applying the method on fMRI datasets collected in a counting Stroop task. It was found that some of the important connections which reveal the interference effects in the counting Stroop task are not present in the results from the patient group. This reflects the conclusion made by the previous literature [26], that females with fragile X have anomalous brain activation during cognitive interference processing tasks and may fail to appropriately recruit and modulate lateral prefrontal cortex and parietal resources.

- The present method was extended to parameter estimation for brain lesion study. The values of conditional probability densities (CPD) reveals how much a region's activation depends on the others. A disconnection or degeneration of connectivity leads to brain disorders. Promise of the present method of studying brain disorders is illustrated by a case study of stroke patient.

## 1.3 Organization of the Thesis

This report is organized so as to describe the research from basic background to specific topic details. It consists of 5 chapters.

Chapter 1 begins with a brief introduction of related methods involved in this project, followed by the motivation and objectives of my research, with an outline of my achieved work.

Chapter 2 gives a review of the framework involved in this project, including basic knowledge about the brain, functional imaging, activation detection, neural system modeling, and earlier approaches for the analysis of brain connectivity.

Chapter 3 introduces Bayesian networks for analysis of functional MR images. The method is tested and compared with earlier methods by using synthetic data. Further, the approach is demonstrated by fMRI data collected in silent word reading task.

Chapter 4 describes the application of the present method on fMRI data collected in a letter search task to introduce a plausible model explaining how activated regions are connected at the levels of three subcomponents of language processing.

Chapter 5 demonstrates how the networks derived from patient and healthy participants by using the present method could be used to explore disconnectivity hypothesis in brain diseases.

Finally, the last chapter concludes this report and gives an outline of my future research plan.

# Chapter 2

# Review of Related Work

This chapter reviews basic knowledge about the brain, functional imaging, activation detection, and neural system modeling. Earlier approaches of effective connectivity analysis including Structural Equation Modeling (SEM), Granger Causality Mapping (GCM), and Dynamic Causal Modeling (DCM) are briefly discussed and their limitations are highlighted.

## 2.1 Background

This section gives some basics related to brain function and fMRI technique, which serves as the threshold leading to the area from where this research started.

### 2.1.1 Basics About the Brain

The brain is composed of three types of tissue: gray matter, white matter, and cerebrospinal fluid (CSF) [27]. Gray matter is called "gray"because it looks relatively dark in anatomical brain specimens (postmortem tissue). Its dark color is produced by densely packed cell bodies of nerve cells (neurons), which perform the basic "command

functions″within our brains. White matter gets its name from the fatty covering of the fibers, which give it a whitish appearance. Neurons are connected to one another by long "wires″that project out of them, called axons which permit individual neurons to send messages back and forth to one another. The brain is bathed inside and out by CSF, a fluid that contains nutrients and byproducts of brain activity. The regions inside the brain that contain CSF are called ventricles. Generally, white matter can be understood as the parts of the brain and spinal cord responsible for information transmission; whereas, grey matter is mainly responsible for information processing. In this research, the interactions among activated brain regions during particular cognitive tasks are studied. It is the relationship between individual information processing that is of interest. Hence, only the grey matter will be considered.

The brain contains 98 percent of body's neural tissue. An average brain weighs about 3 lbs. Ninety five percent of brain's energy is supplied from glucose; fat cannot cross blood-brain barrier [28]. The major regions of the brain include: brainstem, diencephalon, cerebellum, and cerebrum. The cerebrum (also called the telencephalon) is mainly concerned in function imaging analysis, and can be broadly classified into the following:

- **Cortex**: outer surface of cerebrum consisting of gray matter (cell bodies of neurons).
- **Cerebral white matter**: contains neuron axons, and has nuclei of gray matter.

The cortex has expanded over the evolution of the brain; it comprises of 80 percents of the human brain. The cortex consists of two nearly symmetrical hemispheres, left and right, separated by the medial longitudinal fissure. Each hemisphere is subdivided into four lobes: frontal lobe, temporal lobe, parietal lobe, and occipital lobe.

The cortex is highly organized, each of the lobes is thought to be associated with a particular sensory function: frontal lobe for motor functions, temporal lobe for body senses, parietal lobe for auditory function, and occipital lobe for visual functions. Nevertheless, these areas are involved in other functions as well. There are some higher-level association areas that mediate complex functions such as language, planning, memory, and attention.

The complexity of the brain's structure makes it difficult to relate its components to individual functions. One of the objectives of human brain research is to establish structure-function relationships. The general principles of hemispheric and regional brain function are well established but the ways in which specific perceptual and cognitive functions relate to brain structure remain elusive. For details about the brain, please refer to Andreasen (2001) [27] and Frackowiak (1997) [29].

### 2.1.2 Functional MRI

The brain has a grey and rather uniform appearance. However, with the use of Magnetic Resonance Imaging (MRI), a rich internal structure can be visualized. MRI can provide a detailed picture of the brain with a resolution of less than 0.5 mm. Even structures buried deep within the center of the brain can immediately be made apparent. The rich network of arteries, including those penetrating deep into the substance of the brain that are as small in diameter as a hair, can also be seen with MRI [30]. Functional MRI (fMRI) has enabled scientists to look, for the first time, into the human brain in living organism, to literally "watch it while it works". This has revealed exciting insights into the spatial and temporal changes underlying a broad range of brain functions, such as how we see, feel, move, understand each other, and

lay down memories. The technique is safe, allowing repeated examinations to probe time-dependent changes such as those involved in learning. FMRI also improved our understanding of a variety of brain pathologies [30].

MRI of human brain depends on a magnetic property of hydrogen nucleus, known as the spin angular momentum. Like other tissues in the body, over 70 percent of brain is composed of water. Thus, finding out how the water is distributed in the head is just like localizing brain tissue. Different parts of the brain have slightly different amounts of water. Nerve cells, for example, are relatively rich in water, whereas the fatty coating around the long nerve fibres and cells, called myelin, has less water. This generates a contrast between the surface cortex and the underlying white matter of the brain, that can be used to provide exquisite details of brain structure by the technique of MRI [30, 31].

Functional MRI technique has been successfully employed to locate brain processes in various functional experiments involving sensory and cognitive stimuli. Input stimuli in fMRI experiments activate neuronal populations at specific areas of the brain. The metabolic events ensuing neuronal activation create hemodynamic changes such as blood flow and blood oxygenation in the vicinity of activated regions. As de-oxygenated hemoglobin and oxygenated hemoglobin differ in their magnetic properties, neuronal events create magnetic field inhomogeneities in the surroundings. Changes in the magnetic fields in the regions of neuronal activity cause changes in $T_2^*$ (observed signal decay rate, one of the fundamental temporal parameters that used to describe the MR signal [30]) weighted magnetic resonance sequences, and thus result in intensity changes in functional MR images. The contrast mechanism, which depends on the blood oxygenation level, is known as blood oxygen level dependent

(BOLD) contrast.

### 2.1.3 Paradigm Design

The most commonly used paradigm is the regular epochs of stimulus and rest, usually called "on"and "off". The duration of these epochs needs to be long enough to accommodate the hemodynamic response, and so a value of at least 8 seconds, or more commonly 16 seconds is chosen. These epochs are repeated for as long as is necessary to gain enough contrast to noise in order to detect the activation response. The total experimental duration however must be a balance between how long the subject can comfortably lie still without moving, and the number of data points required to obtain enough contrast to noise. There are often some technical limitations to the experimental duration, and there is the possibility of the subject habituating to the stimulus causing the BOLD contrast to reduce with time.

Instead of epochs of stimuli, it is possible to use single events as a stimulus. Again due to the hemodynamic response, these must be separated by a sufficient period of time. This type of stimulus presentation has the major advantage of being able to separate out the relative timings of activations in different areas of the brain. One of the major disadvantages of single event paradigms is that the experiments need to be much longer than their epoch based counterparts, in order to gain the necessary contrast to noise.

### 2.1.4 Correlation Analysis

The technique of correlation analysis exploits a priori knowledge of the anticipate time course of task-related changes in the signal to determine their intensity and spatial

extent [32]. Many current fMRI experiments use a block design in which the subject is instructed to perform experimental (E) and control (C) tasks in an alternating sequence of 20-40s blocks (e.g., CECECEC...). A reference function is constructed by convolving a square wave matching time course of the experimental control task blocks with a fixed model of the hemodynamic response function [33] (an estimate of the fMRI signal changes evoked by a brief burst of neural activity). This reference function is then correlated with the time-series recorded from each voxel as follows:

$$cor(k) = \frac{m_k^T y}{\sqrt{m_k^T m_k}\sqrt{y^T y}} \tag{2.1.1}$$

where $cor(k)$ is the correlation coefficient for the $k$th voxel time-series, $m_k = (m_{k1} \quad m_{k2} \quad \dots \quad m_{kn})^T$ where $m_{ki}$ is the mean-corrected intensity value of the $k$th voxel in the $i$th scan, $y = (y_1 \quad y_2 \quad \dots \quad y_n)^T$ where $y_i$ is the mean-corrected value of reference function at the time point corresponding to the $i$th scan. Those voxels, whose signals are positively correlated with the reference function above a preselected threshold are designated "areas of activation".

### 2.1.5 Statistical Parametric Mapping

Statistical Parametric Mapping (SPM) is a hypothesis-driven method for fMRI analysis. It refers to the construction and assessment of spatially extended statistical processes used to test hypotheses about functional imaging data. These ideas have been instantiated in software that is also called SPM by Wellcome department of imaging neuroscience [34]. It is a voxel-based method by the conjoint use of General Linear Model (GLM) [4] and Gaussian Random Field (GRF) theory [10].

### Spatial Preprocessing

For fMRI time-series analysis in SPM, in order to analyze structural and functional data from different scans of the same subject, or data from different subjects, spatial preprocessing is a necessary step before statistical analysis to make images conform to the same anatomical frame of reference. Spatial preprocessing normally involves: realignment and unwarping, co-registration, normalization and smoothing.

**Realignment and Unwarping** is used in SPM for motion correction. Under the assumption of no shape change and rigid body motion, this intra-modality, intra-subject image registration uses a least squares approach and a 3D (rigid body) spatial translation. The first image in a time-series is used as a reference scan to which all subsequent scans are aligned. These estimates can be used to determine group or subject related movement differences [35]. After realignment, there may still be some motion related artifacts that cannot be solved by rigid transformation [36, 37]. Furthermore, different material has its own susceptibility, so different field disturbance is ensued at different location by the presence of object in the field. This is called susceptibility-by-movement artifacts in fMRI time-series, and unwarping is required in addition to realignment to solve this problem [38].

**Co-registration** often refers to spatial transformation within subject and between image modalities. In fMRI study, it is a necessary step to co-register a series of fMRI images of certain subject to the corresponding structure MR image because only when structural and functional images are co-registered, the high-resolution structural image can be superimposed on fMRI to give functional activation maps of the brain with better anatomical background. Moreover, if functional and anatomical images are registered, the relationship between them for any individuals can be studied [39].

**Normalization** spatially normalizes images into a standard space defined by some ideal model or template image. Generally these algorithms work by minimizing the sum of squares differences between analyzing image and the template. Spatial normalization is performed with affine (12 parameter) and quadratic (3rd order) automated transformations [35, 40].

**Smoothing** convolves spatial filtering of images with an isotropic Gaussian kernel. There are two purposes for this. The first is to make sure that the image data have the characteristics of a random Gaussian field in order that the statistical assumptions of SPM are valid. The second purpose is to recognize that the spatial normalization step is not perfect, and to correct for modest imprecisions in this process by smoothing the data so nearby voxels share more information.

### Statistical Analysis

SPM is a spatially extended statistical process that is used to test hypotheses about regionally specific effects in neuroimaging data [34]. The most established sorts of statistical parametric maps are based on linear models and t-tests.

The general linear model for a response variable $x_{ij}$ (such as the regional cerebral blood flow - rCBF) at voxel $j$ is

$$x_{ij} = \beta_{1j} g_{i1} + \beta_{2j} g_{i2} + \ldots + \beta_{Kj} g_{iK} + e_{ij} \tag{2.1.2}$$

where $i = 1, \ldots, I$ indexes the observation (e.g. scan). The general linear model assumes the errors ($e_{ij}$) are independent and identically distributed normally [$N(0, \sigma_j^2)$]. Here the $\beta_{kj}$ are $K$ unknown parameters for each voxel $j$. The coefficient $g_{ik}$ are explanatory variables relating to the conditions under which the observation $i$ is made.

The model can be rephrased in terms of matrices:

$$x_j = G\beta_j + e_j \tag{2.1.3}$$

Here $x_j = (x_{1j}, x_{2j}, \ldots, x_{Ij})^T$ is the rCBF data vector for voxel $j$, $G$ is the design matrix, comprised of the coefficient $g_{ik}$, with each row for every scan and each column for every effect (condition) in the model. $\beta_j = (\beta_{1j}, \beta_{2j}, \ldots, \beta_{Kj})^T$ is the regression parameter vector for voxel $j$ and $e_j$ is a vector of normally distributed error terms.

The least square estimate of $\beta_j$, say $\widehat{\beta}_j$, is given by:

$$\widehat{\beta}_j = (G^T G)^{-1} G^T X \tag{2.1.4}$$

and the residual error is given by:

$$e_j = x_j - G\widehat{\beta}_j \tag{2.1.5}$$

The hypothesis can be tested to see whether there will be a linear relation linking effect $k$ with the fMRI intensity values of voxels. This can be tested against the null hypothesis that there is no relationship between effect $k$ and the voxel data. On the null hypothesis, $\beta_{kj}$ will not be significantly different from zero. This can be tested by computing a $t$ statistic:

$$t = \frac{\widehat{\beta}_{kj}}{\varepsilon_{kj}} \tag{2.1.6}$$

The standard error $\varepsilon_{kj}$ can be computed from the original analysis of variance, using the remaining error $e_j$ above. SPM performs the above calculation at every voxel in the brain, and thus there is a separate $\widehat{\beta}_{kj}$. This t-statistic is large and positive if the linearity is significantly greater than 0, and large and negative if the linearity is significantly less than 0.

After calculating the $t$-statistic, SPM technique converts the $t$-statistics to $Z$-scores. SPM uses $Z$-scores to display and analyze the $p$-values from the $t$-statistics, which are the numbers from the unit normal distribution that would give the same $p$ value as the $t$ statistic. Then, SPM will show a picture of the $Z$-statistics, thresholded at a given $p$-value.

## 2.2 Connectivity Analysis

A neural system is unique in the sense that it is composed of numerous interconnected elements ranging from single neurons to entire ensembles. These connections range from local intra-regional connections among neurons, to interregional connections among ensembles of neurons across brain areas. Within small and localized regions of the brain, neurons form characteristic sets of connections so-called local circuits. The communication between nerve cells is carried out along physical connections, often linking cells that are separated by large distances. Signals within these connections consist of series of action potentials (spikes) of unit magnitude and duration. The arrival of an action potential at a synaptic junction triggers numerous biochemical and biophysical processes, ultimately resulting in transmission of electrical signals to the postsynaptic (receiving) cells. Neurons in the cerebral cortex maintain thousands of input and output connections with other neurons, forming a dense network of connectivity spanning the entire thalamocortical system.

Brain networks are not random but form highly specific patterns. A predominant feature of brain networks is that neurons tend to connect predominantly with other neurons in local groups. At this high level of scale, connection patterns formed by these local, intra-areal networks are thought to be responsible for the specific

processing requirements of each area. Considering the entire brain, the large-scale organization of the cortex is characterized by patterns of interconnections linking brain areas within and between specific sensory and motor systems.

It is worth noting that communications between and along neural elements underlie the brain function. It was suggested that the brain function is due to the changes in the covariance among neural elements [1]. Looking from higher-order perspective, attempts have been made to demonstrate the contribution of a collection of synchronized brain units to understand brain activities beyond the second-order statistical sense.

As the definitions of brain connectivity are going to be developed, it is useful at this point to introduce and clarify some conventional and new terminology and investigate into their relations. Hitherto, basically three kinds of connectivities have drawn the attentions of scientists and researchers in brain studies: anatomical, effective and functional connectivity. In summary, the definitions are given as follows [41, 42]:

**Anatomical connectivity** simply refers to the set of physical or structural connections linking neuronal units at a given time [28]. In any structural analysis of neural connection patterns, a choice has to be made on the level of the spatial scale at which the analysis is to be performed. Analysis carried out at the local circuit level would most likely focus on the pattern of synaptic connections between individual neurons. Analysis of intra-areal patterns of connections would involve *connection bundles* or *synaptic patches* linking local neuronal populations (neuronal groups or columns). Analysis of large-scale connection patterns would focus on connection pathways linking segregated areas of the brain. Such pathways would comprise of many thousands or millions of individual fibers [43].

**Functional connectivity** refers to the pattern of temporal correlations (or, more generally, deviations from statistical independence) that exist between distinct neuronal units [41, 42]. Such temporal correlations are often due to neuronal interactions along anatomical or structural connections; in some cases, observed correlations may be due to common inputs from an external neuronal source or stimulus source. Deviations from statistical independence between neuronal elements are commonly captured in a covariance matrix (or a correlation matrix), which, under certain statistical assumptions, may be viewed as a representation of the system's functional connectivity. While temporal correlations are perhaps most often used to represent statistical patterns in neuronal networks, other measures such as spectral coherence or consistency in relative phase relationships [44] may also serve as indicators of functional connectivity. The relationship between structural and functional dimensions of brain connectivity is mutual and reciprocal. It is easy to see that structural connectivity is a major constraint on the kinds of patterns of functional connectivity. On the other hand, functional interactions can contribute to the shaping of the underlying anatomical substrate. This is accomplished either directly through activity (covariance)-dependent synaptic modification, or, over longer time scales, through effects of functional connectivity on an organism's perceptual, cognitive or behavioral capabilities, which in turn affect adaptation and survival. The reciprocity between anatomical and functional networks deserves emphasis as it captures some of the unique aspects of brain networks [45].

**Effective connectivity**: Over the past decade, functional neuroimaging has been extremely successful in establishing functional segregation as a principle of organization in the human brain. Functional segregation is usually inferred by the presence of

activation foci in statistical parametric maps [46]. The nature of the functional specialization is then attributed to the sensorimotor or cognitive process that has been manipulated experimentally. Many approaches have been attempted for the connectivity analysis [1, 47, 41, 42, 46, 48, 49, 50, 51]. These approaches have introduced a number of concepts into neuroimaging (e.g. functional and effective connectivity, eigenimages, spatial modes, information theory, multidimensional scaling) and their application to issues in imaging neuroscience (e.g. functional systems, cortical integration, associative plasticity, and nonlinear cortical interactions) [46]. The study of effective connectivity is also gaining importance in relating cognitive theories (eg. attention) to brain operations. It would be very compelling to demonstrate the modulation of responsiveness of these areas in terms of attention-dependent changes in the interactions or effective connectivity among them [50].

Effective connectivity is closer to the intuitive notion of a connection than functional connectivity and can be defined as the influence one neural system exerts over another [41], either at a synaptic or cortical level [46]. Effective connectivity provides insight about causal influence among activated brain regions, while functional connectivity indicates their temporal correlations without causal influence. Effective connectivity gives an accurate quantification to describe the connectivity between brain regions, which could be further divided into direct and indirect effects through the anatomical model [1]. In some aspects, the functional model is close to the notion of effective connectivity since they both measure the dependence between brain regions. Yet, effective connectivity deals with more of a cause-to-effect influence. While, functional connectivity does not necessarily imply a causal link whereas effective connectivity does. To this extent, effective connectivity is much closer to achieving the

goal of investigating and modeling the interactions of neural networks. Thus, the interdependence and reciprocity among these three connectivities deserves emphasis as it captures some of the unique aspects of brain networks [42].

## 2.3 Previous approaches to brain connectivity

This section reviews the earlier approaches of functional and effective connectivity analyses. The studies of functional connectivity are briefly mentioned. The illustration has been focused on more related effective connectivity methods including SEM, DCM, and GCM, followed by a motivation to the present research.

### 2.3.1 Functional connectivity approaches

The study of functional connectivity is to establish connectivity between two regions of the brain on the basis of similar functional response. If two regions of the brain show similar fMRI measurements over time, then they could be determined as being functionally connected, even though there may be no direct neuronal connection between these two regions [52]. There are generally three common approaches for detecting functional connectivity of image data, which are listed as below:

- The correlation based approach is the most direct method that calculates a correlation of image values between pairs of voxels; the correlations are then threshold to reveal the statistically significant connectivity [53].

- Singular value decomposition (SVD) is equivalent to Principal Components Analysis (PCA) [54], and similar to Independent Components Analysis (ICA)

[55]. SVD seeks to express the correlation structure by a small number of principal components, multiplied by random weights that vary randomly over time or subject.

- The clustering method attempts to form clusters of voxels whose values over time or over subject are similar [56]. This is closely related to the correlation method, thresholding correlations, since correlation is used by many clustering methods as a measure of similarity, and thresholding correlations simply clusters together all voxels whose similarity exceeds a threshold value.

Besides, SEM was also used by Goncalves and Hall [57] to model the functional connectivity, but only feasible for connectivity between a small number of pre-selected voxels or regions. More recently, Dodel *et al.* [58] extends conventional activation analyses to the notion of networks by using correlation based method, emphasizing functional interactions between regions independently of whether or not they are activated. Salvador *et al.* [59] explored properties of whole brain networks based on multivariate spectral analysis of fMRI time-series measured in the resting state. Undirected graphs representing conditional independence between multivariate time-series used in the method, was found to be more readily approached in the frequency domain than the time domain.

### 2.3.2 Effective connectivity approaches

The previous methods of effective connectivity including SEM, DCM, and GCM are illustrated and compared with the proposed method in this section.

### Structural Equation Modeling (SEM)

SEM basically represents the covariance structure of a set of variables (as seen in figure 2.1). Instead of minimizing functions of observed and predicted individual values, the SEM minimizes the difference between the sample covariances and the variances predicted by the model [60]; that is, the observed covariances minus the predicted covariances form the residuals. The fundamental hypothesis for these structural equation procedures is that the covariance matrix of the observed variables is a function of a set of parameters. $\Sigma = \Sigma(\theta)$ where, $\Sigma$ is the population covariance matrix of observed variables, $\theta$ is a vector that contains the model parameters, and $\Sigma(\theta)$ represents the covariance matrix predicted by the model.

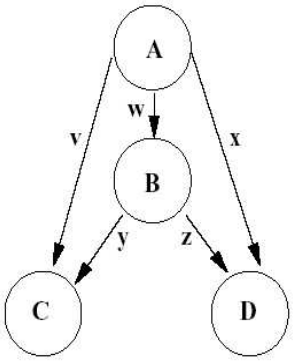
In modeling fMRI, theoretically anticipated connections among $p$ brain regions are written in the matrix form of a path model, and the numerical values of the $q$ nonzero path coefficients in the matrix are estimated so as to minimize a measure of discrepancy between the observed interregional correlation matrix $C$ and the correlation matrix predicted by the model $\Sigma(\theta)$ [47].

The pairwise correlations $c_{ij}$ between the first eigentimeseries for the $i$th and $j$th regions constituted the interregional correlation matrix $C$. The residual variance for each region is estimated by the ratio between the first eigenvalue and the sum of eigenvalues:

$$\Psi_i = 1 - \frac{{\lambda_1}^2}{\sum_{j=1}^{m} {\lambda_j}^2} \tag{2.3.1}$$

where $\lambda_i$ denotes the $i$th eigen value of the correlation matrix.

A model can be written down as a set of simultaneous regression equations:




(a)

**Correlation Matrix**

| | A | B | C | D |
|---|---|---|---|---|
| A | - | | | |
| B | rAB | - | | |
| C | rAC | rBC | - | |
| D | rAD | rBD | rCD | - |

(b)

**Path Equations**

rAB=w
rAC=v+wy
rAD=x+wz
rBC=y+wv
rBD=z+wx
rCD=yz+vx+vwz+ywz+ywx

(c)

**Structural Equations**

$A=\Psi_A$
$B=wA+\Psi_B$
$C=vA+yB+\Psi_C$
$D=xA+zB+\Psi_D$

(d)

Figure 2.1: Schematic representation of methods involved in structural equation modeling of a neural system: (a) Path diagram of a simple network with four regions; (b) The information about the correlations of activities; (c) Path equations showing how the correlations between regions can be decomposed to solve for the path coefficients; (d) Structural equations showing the variance in activity of each region as a function of the weighted variance of other brain regions and a residual influence. [1]

$$x = Kx + \psi \tag{2.3.2}$$

where $x$ denotes the vector of regional variances, $\psi$ denotes the vector of residual variances, and $K$ denotes the path model matrix. Equation 2.3.2 can be rearranged to show that:

$$x = (1 - K^{-1})\psi$$

and the correlation matrix predicted by the path model $\Sigma(\theta)$ is given by:

$$\Sigma(\theta) = (1 - K^{-1})\psi\psi'(1 - K^{-1})'$$

The objective is to find estimates for the non-zero path coefficients $\theta_j$, which minimize a measure of discrepancy between the observed correlation matrix $C$ and the correlation matrix predicted by the model $\Sigma(\theta)$. Estimates of the path coefficients are found by iteratively minimizing the maximum likelihood (ML) discrepancy function:

$$F = \log|\Sigma(\theta)| + \mathtt{tr}(C\Sigma^{-1}(\theta)) - \log|C| - q \tag{2.3.3}$$

where $|.|$ denotes the determinant and $\mathtt{tr}$ denotes the trace of a matrix.

**Granger Causality Mapping (GCM)**

The method proposes to treat a time-series $x[n]$ as a vector autoregressive (VAR) process, and thus to use vector autoregressive models to make inferences on effective connectivity [2]. An inherent problem in inferring interactions at a neuronal population level from fMRI data is the indirect access to the signals of interest (figure 2.2). The fMRI signal can be considered a filtered and sampled version of the Local Field Potential (LFP) signal that is itself a measure of the activity fluctuations of local population of neurons [61].

Taking the temporal structure of signal time-courses into account is related to the commonsense concept of causality: causes always precede effects. Something in the future cannot cause something in the past or present. All events taking place

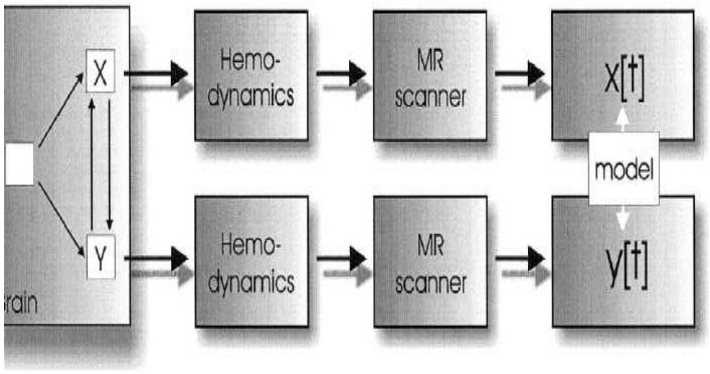



Figure 2.2: Illustration of the indirect access to interacting brain regions with fMRI. Hemodynamics and the MR scanner contribute unwanted artifacts to the signals of interest and might confound modeling efforts. Confounding is especially deleterious when the unwanted contributions are different for the brain regions under investigation (e.g., different hemodynamic responses in different regions) [2].

at a certain point in time must have had their cause at an earlier stage. These considerations have led the econometrist Clive Granger to propose a definition of causality for temporally structured data, i.e., time-series. Conceptually, it amounts to the following: if a time-series $y$ causes (or has an influence on) $x$, then knowledge of $y$ should help predict future values of $x$. Thus, causality (or influence) is framed in terms of predictability.

The vector time-series $x[n]$ can be modeled as an AR process:

$$x[n] = -\sum_{i=1}^{p} A[i]x[n-i] + u[n] \tag{2.3.4}$$

where $u[n]$ is (multivariate) white noise. The matrices $A[i]$ are called the autoregression (AR) coefficients because they regress $x[n]$ onto its own past. $p$ is the order of the auto-regression and will refer to the above model, with adjustable parameters

$A[i]$ and $\Sigma$ to be estimated, as a VAR($p$) model.

Geweke [62] has proposed a measure of linear influence $F_{x,y}$ between the time-series $x[n]$ and $y[n]$ which is regarded as an implementation of the concept of Granger causality in terms of vector autoregressive models. The influence measure $F_{x,y}$ is the sum of three components: the linear influence from $x$ to $y$ ($F_{x\to y}$), the linear influence from $y$ to $x$ ($F_{y\to x}$), and the instantaneous influence between $x$ and $y$ ($F_{x.y}$). The measure can be defined using the residual cross-covariance matrices of the following three VAR models involving the K-dimensional series $x[n]$ and L-dimensional series $y[n]$:

$$x[n] = -\sum_{i=1}^{p} A_x[i]x[n-i] + u[n] \qquad \text{var}(u[n]) = \Sigma_1 \tag{2.3.5}$$

$$y[n] = -\sum_{i=1}^{p} A_y[i]y[n-i] + v[n] \qquad \text{var}(v[n]) = T_1 \tag{2.3.6}$$

$$\text{and with} \qquad q[n] = \begin{bmatrix} x[n] \\ y[n] \end{bmatrix} \tag{2.3.7}$$

$$q[n] = -\sum_{i=1}^{p} A_q[i]q[n-i] + w[n] \qquad \text{var}(w[n]) = Y = \begin{bmatrix} \Sigma_2 & C \\ C^T & T_2 \end{bmatrix} \tag{2.3.8}$$

where $q[n]$ is $O$-dimensional (with $O = K + L$), $\Sigma_1$ and $\Sigma_2$ are $K$ by $K$, $T_1$ and $T_2$ are $L$ by $L$, and Y is $O$ by $O$. It is these residual cross-covariance matrices $\Sigma_1$, $\Sigma_2$, and Y, that are currently of interest, because they quantify how well the current values of x and y from their past valuesto can be predicted (using linear AR models). The measures of total linear dependence between x and y, linear influence from x to y, linear influence from y to x, and instantaneous influence between x and y are defined to be, respectively:

$$F_{x,y} = \ln(|\Sigma_1| \cdot |T_1|/|Y|) \tag{2.3.9}$$

$$F_{x\to y} = \ln(|T_1|/|T_2|) \tag{2.3.10}$$

$$F_{y\to x} = \ln(|\Sigma_1|/|\Sigma_2|) \tag{2.3.11}$$

$$F_{x\cdot y} = \ln(|\Sigma_2| \cdot |T_2|/|Y|) \tag{2.3.12}$$

where "$|\Sigma|$"denotes the determinant of $\Sigma$. From these definitions, it can be seen that it holds that:

$$F_{x,y} = F_{x\to y} + F_{y\to x} + F_{x\cdot y} \tag{2.3.13}$$

**Dynamic Causal Modeling (DCM)**

The general idea behind DCM is to construct a reasonably realistic neuronal model of interacting cortical regions with neurophysiologically meaningful parameters [3]. These parameters are estimated such that the predicted BOLD series, which results from converting the neural dynamics into hemodynamics, correspond as closely as possible to the observed BOLD series. In DCM, neural dynamics in several regions (represented by a neural state vector $z$ with one state per region) are driven by experimentally designed inputs that enter the model in two distinct ways: they can elicit responses through direct influences on specific anatomical nodes (e.g., evoked responses in early sensory cortices) or they can modulate the coupling among nodes (e.g., during learning or attention). DCM models the change in neural states as a non-linear function of the states, the inputs $u$ and neural parameters $\theta^n$:

$$\hat{z} = F(z, u, \theta) \tag{2.3.14}$$

Where $F$ is some nonlinear function describing the neurophysiological influences that activity in all $l$ brain regions $z$ and inputs $u$ exert upon the changes in the others. $\theta$ represents the parameters of the model whose posterior densities are required for inference. It is not necessary to specify the form of equation 2.3.14 because its bilinear approximation provides a natural and useful re-parameterisation in terms of effective connectivity. The bilinear form of 2.3.14 is given as:

$$\hat{z} \approx Az + \sum u_j B^j z + Cu = (A + \sum u_j B^j)z + Cu \tag{2.3.15}$$

$$A = \frac{\partial F}{\partial z} = \frac{\partial \hat{z}}{\partial z} \tag{2.3.16}$$

$$B^j = \frac{\partial^2 F}{\partial x \partial u_j} = \frac{\partial}{\partial u_j} \cdot \frac{\partial \hat{z}}{\partial x} \tag{2.3.17}$$

$$C = \frac{\partial F}{\partial u} \tag{2.3.18}$$

The matrix $A$ represents the effective connectivity among the regions in the absence of modulatory input, the matrices $B^j$ encode the change in effective connectivity induced by the j$th$ input $u_j$, and $C$ embodies the strength of direct influences of inputs on neuronal activity.

DCM combines the above neural model with the biophysical forward model of Friston [63] which describes how neuronal activity translates into a BOLD response (figure 2.3). This enables the parameters and time constants of the neuronal model to be estimated from measured data, using a fully Bayesian approach with empirical priors for the biophysical parameters and conservative shrinkage priors for the coupling parameters. The posterior distributions of the parameter estimates can thus be used to test hypotheses about the size and nature of modeled effects. Usually, these hypotheses concern context-dependent changes in coupling which are represented by

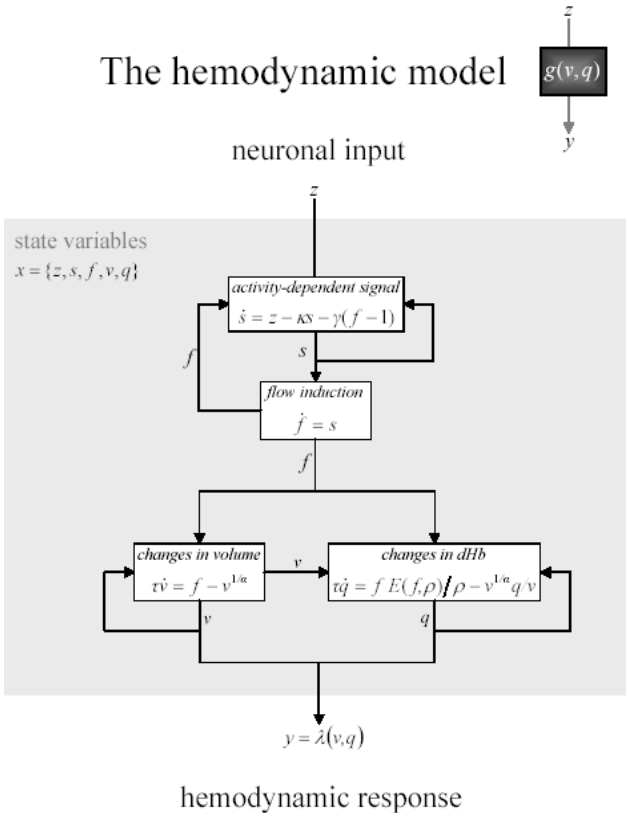



Figure 2.3: The schematic shows the architecture of the hemodynamic model for a single region [3].

the bilinear terms of the model. If there is uncertainty about which connections should be included in a model, or if one would like to compare competing hypotheses (represented by different DCMs), a Bayesian model selection procedure can be used to find the DCM that shows an optimal balance between model fit and the number of parameters.

### Discussion and Motivation

In earlier approaches, the information about neural interactions is often extracted by decomposing interregional covariances among activations. Structural Equation Modeling (SEM) has been the most commonly used method to analyze the effective connectivity among brain regions. McIntosh and Gonzalez-Lima [1] first described SEM and applied for network analysis of vision tasks using PET. Other researchers [51, 64, 65, 66, 67, 68, 69, 70] have later used SEM for the analysis of networks of brain regions involved in sensory or cognitive tasks. In these research, the applications of SEM had been implied to test on several cognitive tasks with known structure in order to confirm or reject a prior hypothesis. Bullmore *et al.* showed how to search for the best fitting covariance model of connectivity from fMRI data by using SEM [47]. Mechelli *et al* [71] constructed a multisubject network based on SEM to illustrate the differences in connectivity among subjects. In this thesis, we propose modeling brain connectivity using Bayesian networks where the interactions are represented by probability distribution functions. Comparing the SEM and the method being proposed, the covariances between the brain regions in SEM limit the behavior of a neural system to the second-order statistical sense whereas the conditional probability densities (CPD) characterising Bayesian network describes the behavior of a network in the complete statistical sense. Hence, complete behavior of a neural system can be described and inferences due to disconnectivity of the network can be made, by using the present method.

Granger causality mapping (GCM), a linear method developed for modelling time-resolved fMR time-series, investigates effective connectivity among activated brain areas by using a vector autoregressive (VAR) model [2]. The connectivity is computed

by evaluating interactions between a current voxel and a reference voxel, and introduce a statistical framework for distinguishing different types of interactions. Moreover, a graphical approach linking the notions of graphical modelling and Granger causality is applied to describe the dynamic dependencies in neural systems [72]. Valdes-Sosa *et al* [73] proposed to use the sparse MAR (SMAR) models to deal with the problem of lacking degree of freedom in some techniques such as multivariate autoregressive (MAR) modelling. Granger causality mapping renders a voxel-wise connectivity analysis whereas the current proposed approach is region-wise and seeks for a global representation of a neural system. The objective of this research is to determine how different regions interact and obtain a holistic view of brain function.

Dynamic Causal Modelling (DCM) was introduced by Friston [3] to model functional interactions at the neuronal level and comprises of a bilinear model for neurodynamics and an extended balloon model for hemodynamics. DCM has shown to be a potential model for making inferences about the temporal changes of effective connectivity from fMRI data [3, 17, 74]. More recently, Penny *et al* [75] proposed a Bilinear Dynamical Systems (BDS) for model-based deconvolution of fMRI time series, which comprise a stochastic bilinear neurodynamical model specified in discrete-time and a set of linear convolution kernels for the hemodynamics. DCM models interactions at the neuronal rather than the hemodynamic level [17], which is more useful in analyzing the temporal interactions between brain regions. Instead, the present approach focuses on exploring the structure of interactions of the neural systems.

Earlier methods of effective connectivity analysis, such as SEM, DCM, and GCM, are confirmatory in the sense that they need a prior connectivity model to begin with. The prior models are often under anatomical constraints and complicated by the fact

that many of them have been obtained in the studies of monkeys. In this thesis and the research group's parallel work [25, 76], the exploratory way of investigating brain connectivity was first proposed. The present method based on Bayesian networks allows not only extraction of the connectivity among brain regions in an exploratory manner, but also modeling of the connectivity pattern in a complete statistical sense.

# Chapter 3

# Effective Connectivity with Bayesian Networks

In this chapter, an exploratory method is proposed to derive effective connectivity from fMRI data, using Bayesian networks. The approach is exploratory in the sense that it does not require an *a priori* model as in the earlier approaches. The conditional probabilities that render interactions among brain regions in Bayesian networks represent the connectivities in the complete statistical sense. The present method is applicable even when the number of regions involved in the cognitive network is large or unknown. The present approach is demonstrated by using synthetic data as well as fMRI data collected in silent word reading task.

## 3.1 Introduction

With the rapid development of medical imaging techniques, researchers are now able to obtain a multifaceted view of brain function and anatomy [77]. Functional brain imaging represents a range of measurement techniques, which extracts quantitative information about physiological function and provides functional maps showing which

regions are specialized for different sensory or cognitive functions [78]. Although many researchers have attempted to identify the individual brain areas involved in various cognitive tasks, holistic views of effective connectivity of higher-order functions have not been investigated thoroughly. More recently, functional integration studies describing how functionally specialized areas interact and how these interactions lead the brain to perform a specific task have become one of the hot topics in brain mapping research [17]. In this chapter, an exploratory approach is presented to determine effective connectivity among brain regions from fMRI data based on Bayesian graphical models where interactions among the regions are represented by conditional probabilities.

The existing methods of connectivity analysis, such as SEM, DCM, and GCM (briefly reviewed in section 2.3), are confirmatory in the sense that they need a prior connectivity model to begin with. The prior models are often under anatomical constraints and complicated by the fact that many of them have been obtained in the studies of monkeys. And it is not always certain which areas are to be included in the study, especially if the brain regions are involved in functions unique to human, such as language and cognition [47]. The present method based on Bayesian networks allows extraction of the connectivity among brain regions from functional MRI data in an exploratory manner. Bayesian network modelling is widely applicable for compactly representing the joint probability distribution over a set of random variables [18]. The nodes represent the activated brain regions and a connection between two regions represents an interaction between them. The Maximum A Posteriori (MAP) estimation of the structure of the functional network is derived from fMRI data to maximize the Bayesian Information Criterion (BIC) by using a greedy search

algorithm.

A synthetic fMRI dataset was used to test the feasibility and robustness of the proposed method. The method was further demonstrated by exploring the functional structure from fMRI data obtained in a silent word reading experiment. The network derived from the silent reading task was consistent with the previous literature and hypotheses, validating the present approach.

## 3.2 Methods

Graphical models (Bayesian networks, Belief networks, probability networks, or causal networks) are models for representing uncertainty in our knowledge [18, 79]. Uncertainty arises in a variety of situations:

- uncertainty in the experts themselves concerning their own knowledge;
- uncertainty inherent in the domain being modeled;
- uncertainty in the knowledge engineer trying to translate the knowledge, and just plain;
- uncertainty as to the accuracy and actual availability of knowledge.

Belief Networks use probability theory to manage uncertainties by explicitly representing the conditional dependencies between the different knowledge components. This provides an intuitive graphical visualization of the knowledge including the interactions among the various sources of uncertainty.

Probabilistic graphical models are graphs in which nodes represent random variable, and the arcs represent conditional dependencies. Hence they provide a compact

representation of joint probability distributions. For example, if we have $N$ binary random variables, a representation of the joint distribution, $P(X_1; \ldots; X_N)$, needs $2^N$ parameters, whereas a graphical model may need exponentially fewer, depending on which conditional assumptions are significant. This aids both inference and learning. There are two main kinds of graphical models: undirected and directed. Undirected graphical models, also known as Markov networks or Markov random fields (MRFs), are more popular with the physics and vision communities. Directed graphical models, also known as Bayesian networks (BNs), belief networks, generative models, causal models, etc. are more popular with the AI and machine learning communities [18]. It is also possible to have a model with both directed and undirected arcs, which is called a chain graph [80]. In this thesis, graphical models are restricted to Bayesian networks.

### 3.2.1 Neural Systems Modeling with Bayesian Networks

A Bayesian network, a specific graphical model that utilizes Bayes' rule for inference, consists of a graph structure and a set of parameters indicating the path coefficients. The graph structure $S$ is a *directed acyclic graph* (DAG) that encodes a set of conditional independence assertions about the variables at nodes. The parameters are represented by conditional probability distributions (CPD) defining the probabilities of the nodes given their parent nodes.

Figure 3.1 shows an example of a Bayesian network, representing a neural system consisting of five brain regions; $\{r_i : i = 1, 2, \ldots 5\}$ denotes the set of brain regions activated during the task where $r_i$ represents the $i$th brain region and $x_i$ denotes the activation of the region; the set of the directed arcs and the conditional probabilities

$\{p(x_i|x_j) : i, j = 1, 2, \ldots 5; i \neq j\}$ characterize the effective connectivity among the brain regions, in the neural system. The brain regions are presumed to collectively and interactively perform the sensory or cognitive task in the fMRI experiment.

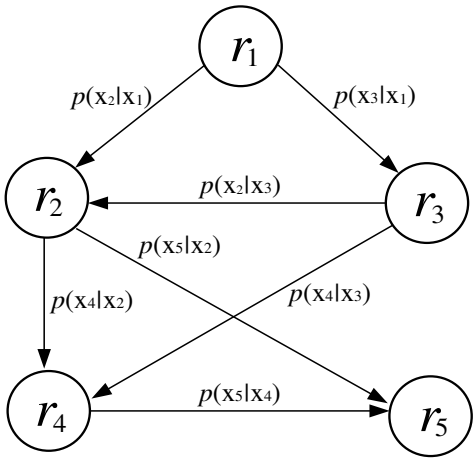



Figure 3.1: Illustration of a neural system represented by a Bayesian network: the set of five activated brain regions $\{r_i : i = 1, 2, \ldots, 5\}$ is represented by the nodes and the set of conditional probabilities, $\{p(x_i|x_j) : i, j = 1, 2, \ldots, 5; i \neq j\}$, represents the interactions.

Consider a neural system consisting of a set of $n$ brain regions $R = \{r_i : \mathrm{i} = 1,2, \ldots, n\}$, that is capable of collectively performing a particular sensory or cognitive task. The activation of a brain region $r_i$ is represented by the average of the time courses of hemodynamic responses of the neurons in the region. Suppose that the average of the time-series responses of the activated brain region is $x_i$. The fMRI experiment is represented by the dataset containing activations of all activated brain regions: $x = \{x_i : i = 1, 2, \ldots n\}$. From the chain rule of probability, the likelihood

of the activation of the neural system is given by

$$p(x) = \prod_{i=1}^{n} p(x_i|x_1, \ldots, x_{i-1}) \tag{3.2.1}$$

where $p(x)$ indicates the joint probability of the activations of all brain regions in the neural system and defines the likelihood of the function of the neural system. For each variable $x_i$, let $a_i \subseteq \{x_j; j = 1, 2, \ldots, n, i \neq j\}$ be a set of parent nodes of $x_i$, that renders $x_i$ and its ancestors conditionally independent. That is,

$$p(x_i|x_1, x_2, \ldots, x_{i-1}) = p(x_i|a_i, \theta_i) \tag{3.2.2}$$

where $\theta_i$ denotes the parameters of the distribution.

Then, a Bayesian network representing the joint probability of the activation of all brain regions, i.e., of the whole brain system, can be written as

$$p(x) = \prod_{i=1}^{n} p(x_i|a_i, \theta_i) \tag{3.2.3}$$

where $\theta_i$ indicates the parameters of the CPD, involving brain region $r_i$ and its parent nodes in $a_i$. Let $\theta = \{\theta_{i,j} : i, j = 1, 2, \ldots n; i \neq j\}$ denote the set of parameters of the whole neural system. It is presumed that all CPD in the graphical model carry the same form.

For two activated regions $r_1$ and $r_2$, the interaction or the influence from region $r_1$ to $r_2$ is indicated by the conditional probability $p(x_2|x_1)$, and the influence from $r_2$ to $r_1$ is $p(x_1|x_2)$. Since the activities of $r_1$ and $r_2$ are not independent, the distribution of $x_1$ will be affected when $x_2$ is given, and vice versa. Thus the interactions of two linked nodes are bi-directional. One of the biggest advantages for choosing Bayesian networks is that they have the bi-directional message passing architecture and can be learned in an unsupervised manner from data.

### 3.2.2 Learning the Structure

Structure learning refers to the learning of the topology of the functional network with respective to the parameterization used. Learning the structure of the neural system from functional MRI data is done by taking a Bayesian approach considering the probability distributions over the parameters or models. This allows the determination of the confidence of one's estimate and the usage of predictive techniques such as Bayesian model averaging [81]. The present model is a directed model, as referred to the Bayesian networks, where all the nodes are fully observed and the interactions are presumed to be Gaussian.

The model attempts to obtain the Maximum a Posteriori (MAP) estimation of the structure, $\hat{S}$, given all the dataset:

$$\hat{S} = \max_{S} p(S|D) \tag{3.2.4}$$

where from Bayes theorem,

$$p(S|D) = \frac{p(D|S)p(S)}{p(D)}; \tag{3.2.5}$$

as the denominator does not depend on $S$, only the numerator is needed to be maximized. $p(S)$ is assumed to have a uniform prior over the structures [82] and to compute $p(D|S)$, the Bayesian approach averages over all possible parameters, weighing each by their posterior probability:

$$p(D|S) = \int p(D|S,\theta)p(\theta|S)d\theta \tag{3.2.6}$$

For large samples, the term $p(D|S,\theta)p(\theta|S)$ is reasonably approximated as a multivariate Gaussian [83]. In addition, approximating the mean of the Gaussian with the maximum likelihood (ML) estimates of $\theta$ and ignoring the terms that do not

depend on the dataset size $N$, The Bayesian Information Criterion (BIC) is obtained to indicate the fitness of the graph to the data:

$$BIC(\theta) = \log\{p(D|\hat{\theta})\} - 0.5l \log\{N\} \tag{3.2.7}$$

where $\hat{\theta}$ is the ML estimate of the parameters and $l$ is the number of free parameters of the model. The present approach assigns a score to each candidate graphical model, which measures how well the graphical model describes the dataset $D$ [84] and yields the best fit model by optimizing the BIC score.

## 3.3 Experiments and Results

In this section, the technique is illustrated with experiments on a synthetic dataset and a fMRI dataset obtained from the fMRI Data Center, Dartmouth College [85]: a silent word reading task (access number: 2-2000-11189). The method was tested on a synthetic dataset for robustness and compared the results with the SEM approach. The structure of the neural system involved in the reading tasks was derived and the validity was investigated with the help of the past literature and known hypotheses.

### 3.3.1 Implementation

In all the experiments, the software package, *Bayes Net Toolbox*, written by Murphy [86] was used for structure learning. The nodes that represent the activated brain regions are set to be *Gaussian*. For each possible structure, a score function describing how well the structure fits the data is calculated. There are two popular choices of the scoring function. The Bayesian score integrates out the parameters, i.e., the marginal

likelihood of the model. The BIC, which is chosen to be used in the present method. has the advantage of not requiring a prior.

There are two different approaches for learning the structure of the network: constraint-based approach and search-and-score approach [18]. The constraint-based approach begins with a fully connected graph and removes edges in a sequential manner if certain conditional independencies are absent in the data. This approach has the disadvantage of repeated independence tests, leading to a loss of statistical power. The more popular search-and-score approach searches through the space of possible DAGs and returns either the best one or a sample of the best models by using a fitness score [86]. Since the number of DAGs is super-exponential of the number of nodes, an exhaustive search in the space is impractical. Either a local search algorithm, such as greedy hill climbing, or a global search algorithm, such as Markov Chain Monte Carlo (MCMC) method [20], should be employed. The Metropolis-Hastings (MH) algorithm [20], an MCMC algorithm, is used in the present method to search the space of DAGs for finding the optimal structure of the network. The sample size is set to be 100 in all the experiments. After learning, the structure which has the highest BIC score is chosen to represent the functional network.

### 3.3.2 Synthetic Data

Synthetic fMRI datasets were generated to test the feasibility and robustness of the proposed method for detecting the underlying neural system.

### Data Generation and Simulation

A neural system was simulated with synthetic time-series where interactions among the brain regions are represented by linear coefficients. Suppose that the activities had zero mean Gaussian variates with a $n \times n$ covariance matrix $\Sigma$, i.e., $N(x; 0, \Sigma)$. Regression equations describe how the activity of one region is related to the activity of the other regions with a set of linear coefficients:

$$x_t = Mx_t + e_t \tag{3.3.1}$$

where $x_t$ denotes the vector of activations of the regions at time $t$ and $e_t$ is the zero mean Gaussian innovation. Matrix $M = \{b_{ij}\}_{n \times n}$ is formed by the predicted interactions among regions. By subtracting $Mx_t$ from both sides of the regression equation and multiplying by $(I - M)^{-1}$, where $I$ is a $n \times n$ identity matrix, the equation becomes:

$$x_t = (I - M)^{-1} e_t \tag{3.3.2}$$

Eq. 3.3.2 can be used to generate synthetic data from a known model given by $M$. The Gaussian variates $e_t$ was randomly generated and then pre-multiplied by $(I - M)^{-1}$. This approach was repeated for each $t$ to obtain the time-series.

All synthetic time-series were simulated to have 300 time points and the dataset was generated based on the following parameters: the structure was the same as in figure 3.1; the nonzero elements of the linear coefficient matrix $M$ were $b_{21} = 1.1, b_{23} = 0.6, b_{31} = 0.8, b_{42} = 1.3, b_{43} = 1.1, b_{52} = 0.9, b_{54} = 1.2$. The present method was used to derive the functional structure from the synthetic dataset.

### Robustness

The synthetic dataset was corrupted by adding random Gaussian noises at randomly selected time points for each time-series to test the robustness of the present method. The percentage of corrupted time points was varied from 10% to 60% in steps of 10%.

A Likelihood-Ratio (LR) measure was used to assess the matching between the learned structure and the known structure as for a given specificity, no other test renders a higher sensitivity [17]. If $p(x|\theta, \hat{S})$ and $p(x|\theta, S)$ are the likelihoods of the estimated structure $\hat{S}$ and the actual structure $S$, then the log of the likelihood ratio is given by

$$\log LR = \log p(x|\theta, \hat{S}) - \log p(x|\theta, S). \tag{3.3.3}$$

Under the null hypothesis that the models are identical, and for large $t$, $-2 \log LR$ is distributed as a $\chi^2$ variable having degrees of freedom equal to the difference in number of parameters between the models. The results of fitness of the present method at various amount of noise are shown in figure 3.2. The values of log $LR$s were scaled between 0 and 1 for better display. The results were stable until 40% of the data points were corrupted by random noise.

### Comparison with SEM

The SEM approach proposed by [47] was used to derive the neural systems generated by the synthetic datasets and the performances were compared with the present technique using Bayesian networks. Several synthetic datasets were generated to simulate brain systems with different number of regions, $n = 3, 4, \ldots, 15$, as illustrated in figure 3.1. The log-likelihood ratios against the number of brain regions are shown in figure 3.3.

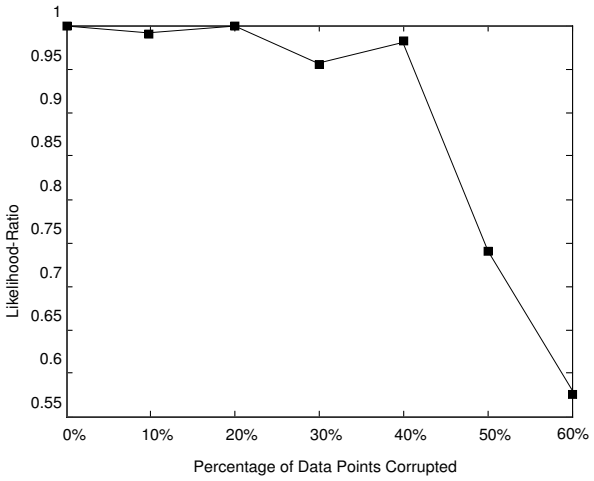



Figure 3.2: Illustration of the robustness of the proposed method for deriving neural systems: the log-likelihood ratios of prediction versus the percentage of number of data points corrupted by random noise

As seen, the present technique using Bayesian networks derived the neural systems closer to the ground truth on all randomly generated synthetic datasets. In the case of synthetic network with 13 regions, the estimated structure did not match well with the actual structure, indicating that the algorithm might have fallen into a local minimum during searching. As the number of regions in the neural system increases, the probability of the structure falling into the local minimum becomes higher.

### 3.3.3 Silent Reading Task

#### Data

The fMRI data used in this experiment consists of six subjects (five males, one female), aged between 20 and 34, with English as the first language. The experiment consisted of a $3 \times 2$ factorial design, three frequencies of presentation: 20, 40, 60 words per

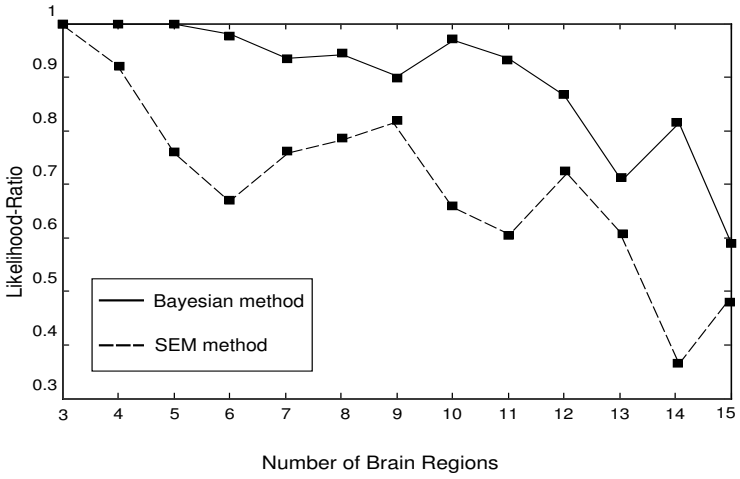



Figure 3.3: Comparison of performances in deriving the functional structure of a neural system by the SEM method and the present approach: the log-likelihoods are shown against the number of brain regions.

minute, and for each, words and pseudowords presentations alternated with a resting condition. The task involved silent reading of words and pseudowords as soon as they appeared on the screen; the resting condition involved fixating to a cross in the middle of the screen. Each subject was presented with 105 words and 105 pseudowords. Stimuli were composed of four, five, or six letters and were presented in 12 blocks. Each block lasted 21s and was followed by a resting period of 16s. The data for each subject contains 360 volume images with a repetition time (TR) of 3.15 s/volume. For more details on the experiment, please refer to [87].

### Detection of Activation

All raw data of functional images were realigned, coregistered, normalized, and smoothed as the preprocessing steps of activation detection. The design matrix, convolved with

a synthetic hemodynamic response function (HRF), was used as the reference waveform for each time-series and then estimated the parameters of the linear model. The time-series were high-pass filtered using a set of discrete Cosine basis functions with a cutoff period of 156s and low-pass filtered using a symmetric HRF as the smoothing kernel to condition the temporal autocorrelations (see [87] for details).

Regions showing increased activity during reading for both words and pseudowords were identified by statistically comparing the fMRI signal while reading relative to the rest condition. The changes in the blood oxygenation level dependent (BOLD) contrast, associated with the performance of the reading task, were assessed on a voxel-by-voxel basis by using the general linear model [4] and the theory of Gaussian fields [10]. This analysis pipeline thus uses multivariate regression analysis and corrects for temporal and spatial autocorrelations of the fMRI data. Group analyses were performed using a fixed-effect analysis (FFX) [88], which infers "typical"characteristics about the group of subjects. Significant hemodynamic changes for each contrast were assessed using the $t$-statistical parametric maps and the results were reported by giving the $t$-values; And the statistical inferences were made at $p < 0.05$ corrected for multiple comparisons by using Family-Wise-Error-Rate (FWER) [89, 90].

SPM2 [4] was used for the above analysis - preprocessing and identification of significantly activated regions. Talairach daemon database [91] and the co-planar stereotaxic atlas [92] were used to assist the specification of the activated regions in Talairach coordinates. The Montreal Neurological Institute (MNI) coordinates given by SPM2 were converted to the corresponding Talairach coordinates by using the technique described by [93]. Table 3.1 and figure 3.4 show the activations found during the silent word reading task. The activations were found in bilateral extrastriate

Table 3.1: Significantly activated brain regions during reading relative to the rest condition are shown in 3D MNI coordinates with $t$-statistics. Statistical inferences were made at $p < 0.05$ corrected for multiple comparisons by using FWER.

| Brain Regions (Broadmann Areas) | Coordinates | $t$-value |
|---|---|---|
| Left Extrastriate Cortex(LEC: BA18,BA19) | (-16 -98 6) | 17.19 |
| Right Extrastriate Cortex(REC: BA18,BA19) | (16 -99 -6) | 17.19 |
| Left Superior Parietal Lobule (LSPL: BA7) | (-28 -60 56) | 7.65 |
| Right Superior Parietal Lobule (RSPL: BA7) | (24 -58 54) | 7.53 |
| Left Middle Temporal Cortex (LMTC: BA21,BA22) | (-50 -52 -8) | 6.51 |
| Right Middle Temporal Cortex (RMTC: BA21,BA22) | (58 -46 8) | 8.13 |
| Left Inferior Frontal Gyrus (LIFG: BA44,BA45) | (-40 12 28) | 7.33 |
| Right Inferior Frontal Gyrus (RIFG: BA44,BA45) | (40 8 30) | 7.46 |
| Left Middle Frontal Gyrus (LMFG: BA46,BA9) | (-48 36 6) | 6.68 |
| Right Middle Frontal Gyrus (RMFG: BA46,BA9) | (40 38 -8) | 6.50 |

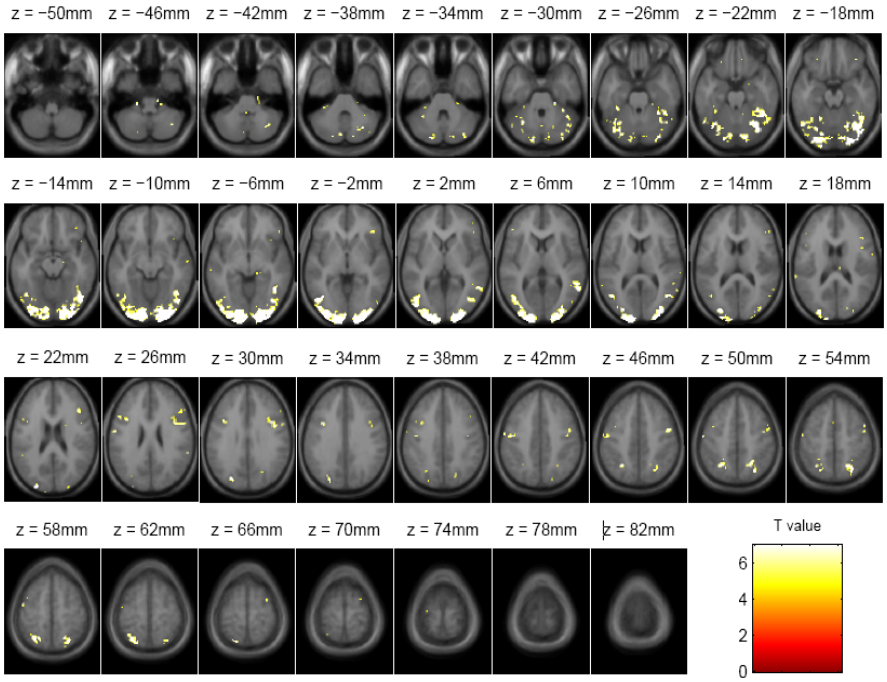



Figure 3.4: Significantly activated brain regions obtained in the group study (using the fixed-effect analysis) of the silent reading task.

cortices, superior parietal lobes, middle temporal cortices, inferior frontal sulci, and middle frontal cortices, and the cerebellum.

### Derivation of Neural System

The time courses of significantly activated regions were extracted by taking the averages of the time-series at the peak-activated voxels and its neighbors (8mm) at the cluster level for all subjects. All extracted time-series representing activated regions were formed into a matrix as the input to learn the structure of the neural system. The Metropolis-Hastings algorithm was used to search the space of all DAGs, with the Bayesian Information Criterion (BIC) as the score function to find the optimal model. Figure 3.5 shows the posterior probability of the DAGs, assuming a uniform structural prior, and each point in the horizontal axis, representing a possible graph structure; the structure with the highest score was chosen to represent the network of this particular task. Figure 3.6 shows the acceptance ratio versus the number of the iteration steps as a crude convergence diagnostic during the search for the optimal structure. The network which had the highest BIC score is shown in figure 3.7.

The left hemisphere has been the focus of the analysis of the neural correlates of reading tasks. Since some language tasks such as those involving different languages, English-knowing bilinguals, literate versus illiterate, etc., show activation in both hemispheres [70, 94, 95], all the activated regions of the cortex were included and the possible connections among all the brain regions were explored.

The extrastriate cortex (EC: BA18, BA19) in the visual cortex plays the role of visual representation in word processing [96]. The connection from the extrastriate cortex to superior parietal lobe (SPL: BA7) forms the dorsal stream of visual analysis,

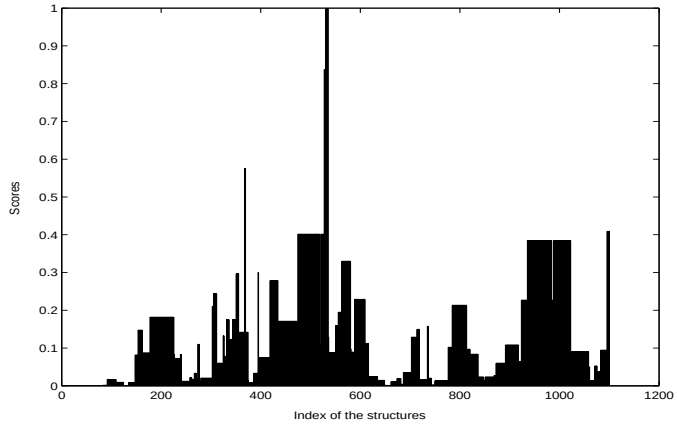



Figure 3.5: The posterior scores of possible DAGs derived from the Metropolis-Hastings algorithm, assuming a uniform prior of the structures.

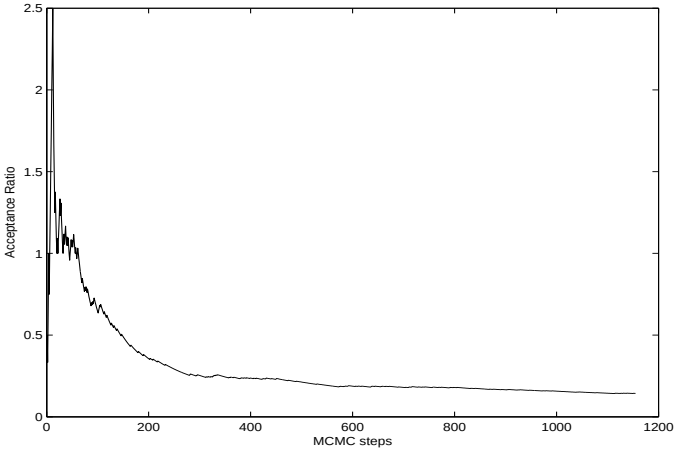



Figure 3.6: The acceptance ratio versus the number of MCMC steps in finding the optimal structure of the neural system.

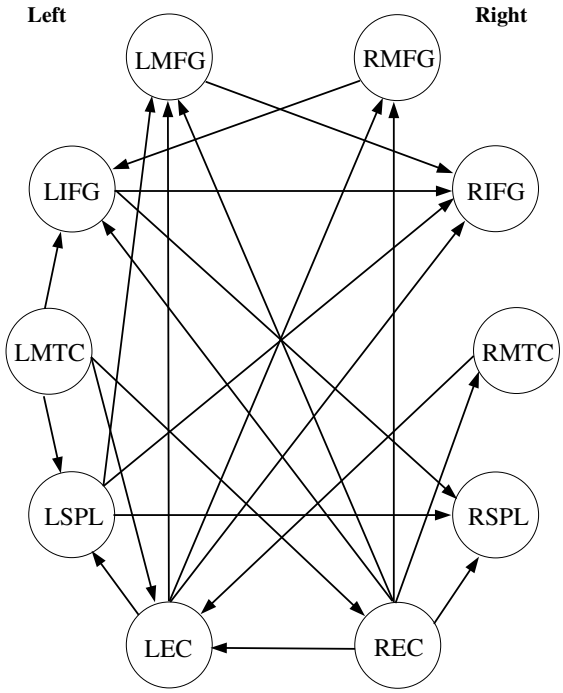



Figure 3.7: The neural system learned from fMRI data of the silent reading task. L(R)EC: left (right) extrastriate cortex, L(R)SPL: left (right) superior parietal lobe, L(R)MTC: left (right) middle temporal cortex, L(R)IFG: left (right) inferior frontal gyrus, L(R)MFG: left (right) middle frontal gyrus.

performing the perception of visual word form. As seen in figure 3.7, the connections from EC to SPL are found in both hemispheres (LEC→LSPL and REC→RSPL). Meanwhile, the connections from the EC to prefrontal cortex including middle frontal gyrus (MFG: BA46, BA9) and inferior frontal gyrus (IFG: BA44, BA45) represent the information flow for the processing of semantic analysis and decision (LEC→LMFG, LEC→RMFG, LEC→RIFG, REC→RMFG, REC→LMFG and REC→LIFG) [47]. Furthermore, the connections between EC and middle temporal cortex (MTC: BA21, BA22), associated with the retaining and recalling of words from the memory [96] are found in both hemisphere with reversed directions (REC→RMTC, LMTC→LEC); the reversed direction may be due to the bi-directional characteristic of the connectivity, represented by the Bayesian networks. In addition, a homologous interhemispheric connection between the ECs of both sides (REC→LEC) is found, which may be due to the transcallosal inferences between two hemispheres [51].

The parietal lobe generally performs the integration of sensory information for the control of movement. In particular, the superior parietal lobe (SPL: BA7) plays the role of visual analysis and mainly makes efferent connections to the prefrontal cortex including MFG and IFG, providing more elaborate information (LSPL→LMFG, LSPL→RIFG) [96]. A homologous interhemispheric connection is also found between the SPLs (LSPL→RSPL). As seen in figure 3.7, the functional links from EC via SPL to prefrontal cortex form the dorsal visual pathway of language processing (LEC → LSPL → LMFG, LEC → LSPL → RIFG) [51].

The temporal lobes are involved in understanding and processing language, intermediate and long term memory, complex memories, the retrieval of language or words, and emotional responses [97]. The middle temporal cortex (MTC: BA21,

BA22) involved in the model is the general association cortex that integrates the input from the lower level auditory and visual areas for retaining in the memory. In particular, the posterior aspect of the left middle temporal cortex, which is also called the Wernicke's area, is involved in storing the visual word forms and processing lexical-semantic information [98]. It is supposed to have connections with LSPL for movement control (LMTC→LSPL), with the prefrontal cortex for semantic phonologic retrieval and semantic processing (LMTC→LIFG) and with EC for memory retention (LMTC→LEC, LMTC→REC) [99, 100, 101].

The MFG is involved in tasks that require executive control, such as the selection of behavior based on short term memory [64]. It receives inputs from the posterior parietal and temporal lobe. The IFG is most active for phonemic decisions and receives inputs from temporal lobes and parietal lobes [97, 99]. As seen in figure 3.7, except for the connections that have been mentioned above, there are interhemispheric connections between the prefrontal regions, including the interconnection between the homologous regions of IFG (LMFG → RIFG, RMFG → LIFG, LIFG → RIFG), which may be involved in semantic processing during inner speech [47].

In the derived interhemispheric language network, the left hemisphere showed dominant pathways, which is consistent with the traditional language network (left-hemispheric). Although the activations have been symmetrically distributed in both hemispheres, the right hemisphere activations may be due to the transcallosal influence of the left. This hypothesis is supported by the fact that there are more connections between the regions in the left hemisphere and the regions in the right hemisphere receive only results of processing in the left regions.

Table 3.2: List of the connections among the activated brain regions, found to be involved in the silent reading task, which had been previously verified in other language-based tasks.

| Connection | Functional Discription | Relative Reference |
|---|---|---|
| LEC → LSPL | perception of visual word form | Horwitz [102] |
| REC → RSPL | perception of visual word form | McIntosh [51] |
| LEC → LMFG | semantic decision and analysis | Krause [64] |
| | | Bullmore [47] |
| REC → RMFG | semantic decision and analysis | Krause [64] |
| REC → LEC | homologous interconnection | McIntosh [1] |
| | | McIntosh [51] |
| | | Krause [64] |
| LSPL → LMFG | executive control | Honey [67] |
| LSPL → RIFG | phonemic decisions | Honey [67] |
| LSPL → RSPL | homologous interconnection | Honey [67] |
| LMTC → LSPL | semantic processing | Price [99] |
| | | Horwitz [102] |
| LMTC → LIFG | semantic phonologic retrieval | McKiernan [69] |
| | | Matsumoto [103] |
| | | Hampson [100] |
| | | Mechelli [71] |
| LMTC → LEC | memory retention | Nyberg [65] |
| LMTC → REC | memory retention | McIntosh [51] |
| LMFG → RIFG | inner speech production | Krause [64] |
| | | Nyberg [65] |
| | | Petersson [70] |
| LIFG → RIFG | homologous interconnection | Honey [67] |

The connections in the derived model that are consistent with the previous literature are given in table 4.2. Due to the fact that the specific networks for each cognitive task are different even though the tasks are very similar (e.g., different presenting rate, different words, or different block design in reading tasks), the existing literature can only be used as a general reference to an existing connection. The connectivity pattern derived from the present method is consistent with the information flow in the silent reading task as evidenced by the literature, but the connections without a corresponding reference cannot be refuted.

## 3.4 Discussion

Earlier approaches to neural system analysis, such as SEM, DCM, and GSM, are confirmatory; a researcher is more likely to use them to determine whether a previously known neural system model is valid, rather than to "find"a suitable model from the data [104]. The structures of those models were constrained by the prior model derived from previous studies or by anatomical constraints although the exact model for the experiment under consideration is often unknown. The present method investigated the use of Bayesian networks to learn large or unexplored cognitive networks from fMRI data by assuming that the basis of such networks does not have proper prior models.

In SEM studies, effective connectivity was explored using path coefficients indicating the covariances among regions [47]. The present approach uses conditional probability densities in graphical models to determine the structure of a functional network. In contrast to the second-order models, such as SEM, the connections

between the regions in the present approach were derived by considering CPD describing the behavior of a network in the complete statistical sense, which renders more information about the effective connectivity. The results on the synthetic data showed that the Bayesian networks can better fit the functional imaging data than the covariance-based models. The connectivity analysis by GSM is voxel-wise; in contrast, the present approach is region-wise and seeks for a global representation of a neural system. Both DCM and the present approach make inferences about the connectivity of the network in the Bayesian framework, therefore there are no limits on the number of connections that can be modelled without an overfitting problem. However, DCM analyzes interaction at the neuronal rather than the hemodynamic level, which is more useful in analyzing the temporal interactions between brain regions. Instead, the present approach focuses on exploring the structure of interactions of the neural systems.

The complexity of the brain makes it difficult to be explored, especially in higher cognitive tasks; the analysis of functional integration (functional connectivity and effective connectivity) is still far from settled. The proposed method of exploring global neural systems from functional imaging data provides an alternate method to study brain function in terms of networks. The network derived from the present method for silent reading was consistent with the literature, providing a partial validation of the present approach, though a gold standard for the networks of the tasks considered is unavailable. In the silent reading task, the network demonstrated the dominance of language processing in the left hemisphere and the regions in the right hemisphere receives the effects of processing from the left hemisphere.

The structure of the present functional brain network was determined from the

data by the present method in a completely exploratory manner. As seen in the experiments with synthetic data, the present method was robust to random noise and outperformed SEM in determining the structure. The MCMC algorithm searches the possible DAG space and returns a sample of structures found after search-and-score learning. The structure with the highest score was chosen as it matches the data the best. This may not always be the best choice because of the local minima problem. The synthetic data simulation shows that as the number of region increase, the search has a higher rate of falling into local minima. However, this problem can be mitigated if a priori knowledge (even if incomplete) of the regions of activation or their connectivity is available. A compromise between confirmatory and exploratory approaches might be more appropriate for brain connectivity analysis.

Another limitation of the present method is that: the strengths of connections in networks derived from our statistical model is within the range [0, 1], which is not available for differentiating between the negative and positive influences. The facilitating and suppressing of the interactions is normally deal with functional connectivity analysis, where the correlations between each pair of regions are studied. The main focus of the present method in this thesis is the structure learning, where only the links with positive significant effect is noted. In the future research, the method could be extended to combine with the correlation method, so that both positive and negative influences can be addressed.

The ability of exploring structure from data by using the present method can be used to introduce possible neural system model based on specific fMRI experiments. This approach had been proposed to explain a language system in Chapter 4. This ability could also be extended to differentiate the performance of patients

and healthy participants performing the same cognitive tasks, and explore disconnectivity hypotheses in brain disease as illustrated in Chapter 5. The main aim of the present work is to determine the existence of significant interactions among brain regions. Estimating the strengths of these interactions and exploring the behavior of such networks due to an abnormal event such as a stroke is described in Chapter 5. Another major advantage of Bayesian networks might be its ability to infer network function in the case of brain disorders, as inferencing is a strength of the graphical models. This feature could be used to mimic the disconnectivity hypothesis in lesion study. Mimicking lesion study is remained as a future research.

# Chapter 4

# Study of Language System

Over the past decade, many new and non-invasive neuroimaging techniques have allowed researchers to investigate specific brain regions involved in language processing. Previous studies investigating the neuroanatomical representation of English at the levels of orthography, phonology, and semantics are reviewed and discussed in this chapter. However, these neuroimaging studies have not identified the interactions among brain regions involved in language tasks. The present method is applied on fMRI data collected in letter searching task to determine a plausible neural system explaining how different regions interact for language processing at the level of three subcomponents: orthography, phonology and semantics.

## 4.1 Introduction

The classic model of language organization was developed in the late 19th century. The basic tenet of this model is that receptive speech is localized to the left posterior temporal Wernicke's area (BA 22) and expressive speech is localized to the left inferior frontal gyrus or Broca's area (BA 45). However, the Wernicke-Broca model, which was human-lesion based (i.e., on the location of strokes that altered language), may

be oversimplified for two reasons. Firstly, many of these strokes affected regions larger than the classic Wenicke's and Broca's area. Secondly, lesions restricted to the classic Wernicke's and Broca's areas do not produce the entire spectrum of symptoms associated with these characteristic aphasias [105]. Although the classic model of language identifies certain critical brain regions for language processing, it does not address the linguistic representation of language. Kandel [106] argues that a more comprehensive approach to the study of language should emphasize language competence and principles of grammar rather than merely focus on language performance.

Past neuroimaging studies have demonstrated that there are differences in the neuroanatomical representation of English at the levels of orthography, phonology, and semantics. Evidence from these studies suggest that the differences lie in the location of the brain regions activated, i.e., orthographic processing is mainly subserved by the occipital regions, phonological processing is subserved by the left fronto-parietal regions and semantic processing by the left fronto-temporal regions. Although fMRI allows us to identify certain brain regions associated with language processing, previous research did not provide enough information regarding the functional relations between the regions involved. As Buchel, Frith and Friston [107] note, fMRI does not reveal how the cognitive processes interact or how the activated brain regions communicate with each other. Brown and Hagoort [108] argue that elucidating both the functional connectivity and interactivity of language-related cortical regions are essential parts of the cognitive neuroscience of language.

As mentioned in section 3.1, the existing methods of connectivity analysis are confirmatory in the sense that they need a prior connectivity model to begin with. In what follows, a neuroanatomical network model of language processing is derived

across the three main components i.e., orthography, phonology and semantics, using the data-driven method. The validation for the network derived are given with the help of reported results from past neuroimaging studies. As it may be difficult to truly isolate any one particular component in language processing, it is worthwhile to discuss the model to the three components collectively. FMRI data was collected in a letter searching task, which was designed to have a cross comparison among reading of high frequency word (HFW), low frequency word (LFW), legal non-word (LNW) and illegal non-word (INW). This will probably reveal the differences between the three components of language processing. The dataset was applied with the present method to derive and discuss a plausible model for language system.

## 4.2 Brief Review of Language System

Language is one of the most precious abilities for human being, yet most of the people take it for granted. It is not realized that how much a human depends on the ability to talk, listen, and read. The use of language is one of the most complex skills, there are many ways to approach the study of language [109]. This section gives a brief introduction of the language system.

### 4.2.1 Component Processes of Language

Although most of people probably think of words as the meaningful units of language, linguists break down language somewhat differently. Cognitive scientists investigating the neuronal correlates of cognition often aim to tease apart the components of a cognitive process. The complex process of reading involves distinct but interrelated

subcomponents, such as (a) orthography - the knowledge of letter combinations in written words; (b) phonology - the knowledge of the sound structure of words; and (c) semantics - the knowledge of the meaning of words. Vocal intonations that can modify the literal meaning of words and sentences are collectively called prosody. The stringing together of sentences to form a meaningful narrative is called discourse. There are analogues in visual language, such as American Sign Language (ASL, or Ameslan). A morpheme in ASL would be the smallest meaningful movement.

### 4.2.2 Localization of Language

Current ideas about the localization of language processes come from four basic lines of inquiry: (1) anatomical studies of language, (2) studies of lesions in human patients, (3) studies of brain stimulation in awake human patients, and (4) brain imaging studies.

#### Anatomical Areas

The anatomical landmarks used by researchers for describing brain regions associated with language vary considerably. Some researchers refer to sulci, others to Brodmann's areas, and still others to areas associated with syndromes, such as Broca's area and Wernicke's area. Language regions include the inferior frontal gyrus and the superior temporal gyrus, in which Broca's areas and Wernicke's areas, respectively, are located. Parts of surrounding gyri, including the ventral parts of the precentral and postcentral gyrus, the supramarginal gyrus, the angular gyrus, and the medial temporal gyrus, also are within the core language regions. Other regions taking part in language include the dorsal part of BA 6 within the lateral fissure (supplementary

motor area); parts of the thalamus, the dorsolateral parts of the caudate nucleus, and the cerebellum; visual areas (reading), sensory pathways, and motor pathways; and pathways connecting all of these various regions. Furthermore, many regions of the right hemisphere also have roles in language.

**Lesion Studies**

Most discussions of the neural basis of language have centered on Broca's area and Wernicke's area. The early neurological model of language by Wernicke, as well as its later revival by Geschwind, now called the Wernicke-Geschwind model, was based entirely on lesion data. This model has played a formative role in directing research and organizing research results. It has the following three parts: (1) The meaning of words is represented in Wernicke's area. When a person listens to speech, word sounds are sent through the auditory pathways to the primary auditory cortex, Heschl's gyrus. From there, they are relayed to Wernicke's area, where the sense of the words is extracted; (2) To speak, it is necessary to send word meanings over the arcuate fasciculus to Broca's area, where morphemes are assembled. The model proposes that Broca's area holds a representation for articulating words. Instructions for speech are sent from Broca's area to the adjacent facial area of the motor cortex, and from there instructions are sent to facial motor neurons in the brainstem, which relay movement commands to facial muscles; and (3) Reading requires that information concerning writing be sent from visual areas BA 17, 18, and 19 to the angular gyrus (BA 39) and from there to Wernicke's area, which reads silently or, in conjunction with Broca's area, reads out loud. Although conceptually useful, many aspects of this model have been modified by improved lesion analysis and by brain imaging studies.

### Electrical Stimulation

The language zones of the neocortex, particularly those pertaining to speech, were identified by Penfield and others by using cortical stimulation during surgery. Statistical analysis of results from hundreds of patients have made it possible to construct a map of these regions. Several important conclusions had been drawn from the experiments: (1) the experimental data do not support strict localizationist models of language because the effects of stimulation of the anterior and posterior speech zones on speech functions are remarkably similar; (2) stimulation of the neocortex considerably beyond the classical areas of Broca and Wernicke disturbs speech functions; and (3) stimulation of the speech zones affects more than just talking because it produces deficits in voluntary motor control of facial musculature as well as in short-term memory and reading.

### Functional Imaging

With the development of PET, fMRI, and event-related potential procedures, cognitive psychologists have become more interested in the neural correlates of language processing. The fMRI studies had been used to identify candidate language processing areas in the intact human brain and to distinguish these from non-language areas [110]. The studies suggest that (1) Wernicke's area, although important for auditory processing, is not the primary location where language comprehension occurs; (2) language comprehension involves several left temporoparietal regions outside Wernicke's area, as well as the left frontal lobe; and (3) the frontal areas involved in language extend well beyond the traditional Broca's area to include much of the lateral and medial prefrontal cortex.

### 4.2.3 Cognitive Disorders Effecting Language

Although the deprivation of any function is onerous, diseases that affect cognition are devastating to humans in a particular way. Not being able to communicate thoughts efficiently can cut a person off from his or her livelihood and family and have immense effects on emotional state and social position. Language can be impaired by sudden events such as stroke or head injuries, insidiously progressive conditions such as Alzheimer's Disease or Parkinson's Disease, or developmental disorders as happens in dyslexia.

We are now able to make highly specific diagnoses of what language processors are affected in a particular language disorder, and recent work has begun to demonstrate that targeting these specific impairments can improve language functioning. As we know more about the brain mechanisms involved, medical therapies such as those that improve attention will also become more tailored to remediation of particular language disorders. The future holds much promise for applying our rapidly-accruing knowledge regarding the neural basis of language to improving the quality of life of language-impaired individuals [111].

### 4.2.4 Right Hemisphere Contributions to Language

Although there is little doubt that the left hemisphere of right-handed people is the dominant hemisphere in language, there is growing evidence that the right hemisphere does have language abilities. The best evidence has come from studies of split-brain patients in whom the linguistic abilities of the right hemisphere have been studied systematically with the use of various techniques for lateralizing input to one hemisphere. The results of these studies have shown that the right hemisphere has little

or no speech but surprisingly good auditory comprehension of language, including both nouns and verbs. There also appears to be some reading but little writing ability in the right hemisphere. In short, although the right hemisphere appears to be able to recognize words (semantic processing), it has virtually no understanding of grammatical rules and sentence structures (syntactical processing).

Complementary evidence of the right hemisphere's limited role in language comes from studies of people who have had the left hemisphere removed, a procedure known as hemispherectomy. If the left hemisphere is lost early in development, the right hemisphere can acquire considerable language abilities although people with left hemispherectomies are by no means normal. Left hemispherectomy in adulthood is far more debilitating, and in all such cases there are severe deficits in speech; but even these people have surprisingly good auditory comprehension. Their reading ability is limited, however, and writing is usually absent. In general, it appears that left hemispherectomy produces language abilities that are reminiscent of those achieved by the right hemisphere of commissurotomy patients.

## 4.3 Experiments and Results

The Bayesian framework is applied on a letter searching fMRI task to derive a plausible model for the neural system of language.

### 4.3.1 Data

Fourteen right-handed volunteers (eight females; 18 to 23 years of age), with no known neurological disorders, volunteered for the imaging study. They gave informed

consent for this study and received a token sum upon completion of the experiment. All participants were screened behaviourally for language abilities using the Language Background Questionnaire and a battery of paper and pencil language subtests. The three criteria for inclusion in this study were that participants: (a) scored at least 70% or better in accuracy for all the language screening tests, as these scores were used as an indication of competence in both languages; (b) be right-handed (based on self-report); and (c) have no known history of neurological impairment.

The visual search task (VST) allows us to use a variety of word stimuli to investigate word processing in the human brain. The VST is used to investigate orthographic processing directly, and phonological and semantic processing indirectly.

Four types of word strings were used as stimuli in the VST, namely:

1. High frequency words (HFW) - e.g., "brain", "sweet";
2. Low frequency words (LFW) - e.g., "dense", "obtuse";
3. Legal non-words (LNW) - e.g., "frell", "drurl";
4. Illegal non-words (INW) - e.g., "sghri", "ddaor"

Imaging was performed on a 1.5 Tesla whole-body magnetic resonance imaging (MRI) scanner (Siemens Vision; Erlangen, Germany) at the Singapore General Hospital. Prior to imaging, participants were briefed on the scanning procedures and experimental conditions so as to minimize anxiety and enhance task performance. Participants were asked to lie supine inside the MRI scanner with their heads inside a standard head coil. Head movement was minimized within the head coil using foam wedges and a restraining band was placed across the forehead. Participants were also

fitted with headphones ($MSI^{TM}$, Tampa, Florida) that attenuated ambient scanner noise by 30dB. They were able to receive instructions using these headphones before each run commenced.

The presentation of written words and characters was controlled by a laptop, located outside the scanning room, running the computer E-prime software. Stimuli were back-projected via a high- resolution LCD projector ($MSI^{TM}$, Tampa, Florida) onto an opaque screen positioned at the head end of the bore. Participants viewed the screen through a specially designed mirror mounted on the head coil. They were asked to respond to the visually presented words and characters by pressing one of two response buttons placed in each hand. For all trials, participants were instructed to indicate a 'yes' response by pressing the button with their first finger and a 'no' response by pressing the button with their second finger i.e., the fMRI setup was made as similar as possible to the behavioural procedure.

Tri-planar scout images in the sagittal plane and T1-weighted 3D coronal anatomical images (MPRAGE sequence) were acquired. These high-resolution images of the entire brain served as structural scans for Talairach transformation. Functional images were then obtained with a T2-weighted gradient echo, echo planar imaging (EPI sequence, 17 contiguous oblique axial 8mm slices, TR/TE/$\theta$ = 3000ms/66.00ms/$90^o$, voxel size = $1.5 \times 1.5 \times 8.0$ mm, FOV = 190 mm, and acquisition matrix = $128 \times 128$) with blood oxygen level-dependent (BOLD) contrast. For each of 17 slices, 360 images were acquired in 1460s with a 1000ms cue i.e., "!", prior to the first trial.

### 4.3.2 Detection of Activation

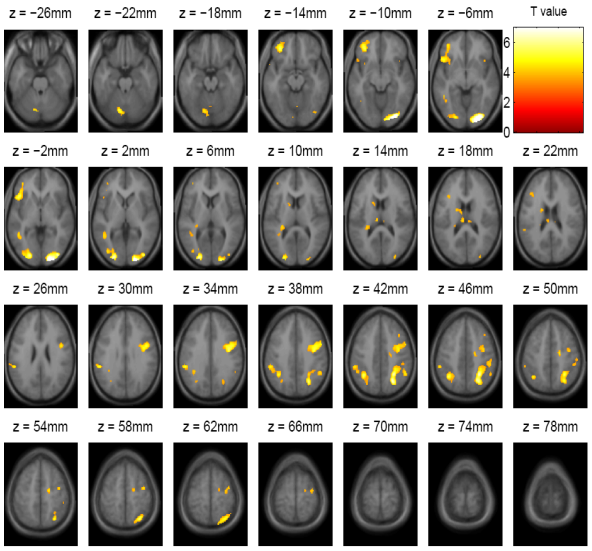
fMRI data were analyzed using SPM2. The first five scans (15s) in each session (during which magnetization steady state was being reached) were excluded from data analysis. Prior to statistical analysis, all functional images were corrected for movement using least-squares minimization [37]. The functional images were then coregistered to the subjects' 3D T1-weighted image. Using the 3D image as a guide, the functional images were then spatially normalized into the SPM standard space. Images were then resampled at every 2mm using Sinc interpolation and smoothed with a FHWM 8mm, 3D Gaussian kernel to decrease spatial noise. Changes in blood oxygenation level dependent (BOLD) contrast associated with the performance of the reading tasks were assessed on a pixel-by-pixel basis, using the general linear model [4] and the theory of Gaussian fields [10] as implemented in SPM2. The method takes advantage of multivariate regression analysis and corrects for temporal and spatial autocorrelations in the fMRI data. Group analyses were investigated using fixed-effect analysis (FFX) [88] which infer "typical"characteristics about this group of subjects. Each language effect was tested by applying appropriate linear contrast. Significant hemodynamic changes for each contrast were assessed using the $t$-statistical parametric maps and results were reported by giving the $t$-values. Activations below a threshold of $p < 0.05$ corrected for multiple comparison using Family-Wise-Error-Rate (FWER) [89, 90] were reported. The ICBM single subject MRI anatomical atlas from the ICBM consortium [112] was used to interpret the locations of the activations. The pre-labeled high resolution (0.5 mm isotropic) atlas was normalized, using the same transformation as the MR images, to infer MNI coordinates and region labels. Talairach coordinates are not given but interested

Table 4.1: Results of group activation for the four tasks, activation voxels are shown in 3D MNI coordinates ($p < 0.05$, FWER corrected). 14 subjects included in the analysis. (EC: Extrastriate cortex; SPL: superior parietal lobe; IPL: inferior parietal lobe; IFG: inferior frontal gyrus; VIFG: ventral inferior frontal gyrus; PMA: primary motor area; SMA: supplementary motor area; FUG: fusiform gyrus; letters L and R before the regions indicating the left and right hemisphere)

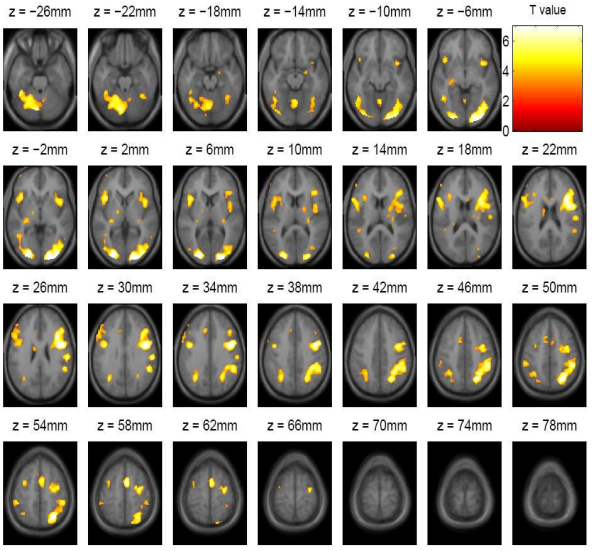
| Brain regions | High Frequency | | Low Frequency | | Legal non-word | | Illegal non-word | |
|---|---|---|---|---|---|---|---|---|
| | Coordinates | $t$-value | Coordinates | $t$-value | Coordinates | $t$-value | Coordinates | $t$-value |
| Occipital | | | | | | | | |
| LEC: BA18,BA19 | -21, -93, 4 | 9.99 | -17, -89, 6 | 9.28 | -19, -85, -4 | 11.51 | -21, -90, 8 | 9.20 |
| REC: BA18,BA19 | 23, -87, 6 | 6.51 | 25, -87, 4 | 10.10 | 25, -81, 2 | 10.13 | 25, -85, 4 | 7.61 |
| Parietal | | | | | | | | |
| LSPL: BA7 | -25, -60, 41 | 6.70 | -27, -57, 48 | 8.61 | -25, -59, 47 | 8.59 | -27, -57, 47 | 10.90 |
| RSPL: BA7 | | | | | 29, -59, 50 | 6.16 | 29, -59, 49 | 6.89 |
| LIPL: BA39,BA40 | -29, -46, 42 | 6.79 | -37, -38, 38 | 7.14 | -37, -42, 40 | 5.77 | | |
| RIPL: BA39,BA40 | 35, -50, 43 | 5.90 | 31, -48, 46 | 5.04 | | | 39, -50, 46 | 7.52 |
| Frontal | | | | | | | | |
| LIFG: BA44,BA45 | -39, -4, 31 | 5.79 | -43, 1, 29 | 9.65 | -43, 1, 31 | 8.23 | -47, 3, 27 | 9.88 |
| RIFG: BA44,BA45 | | | 43, 1, 31 | 6.43 | 43, 3, 27 | 6.87 | 45, 3, 29 | 8.69 |
| RVIFG: BA47 | 43, 23, -2 | 5.67 | 39, 17, 0 | 6.16 | 35, 19, -7 | 4.94 | | |
| LPMA: BA4 | -53, 5, 34 | 4.90 | -31, -8, 55 | 6.52 | -39, -5, 51 | 6.84 | -39, -7, 46 | 6.93 |
| SMA: BA6 | -17, -9, 48 | 4.84 | -3, 2, 55 | 7.25 | -7, 8, 51 | 5.48 | -1, 6, 53 | 7.35 |
| Temporal | | | | | | | | |
| LFUG: BA37 | -43, -49, -16 | 2.78 | -37, -49, -16 | 4.75 | | | | |
| RFUG: BA37 | 39, -35, -15 | 3.01 | 47, -50, -2 | 5.38 | 39, -45, -17 | 4.78 | | |

readers can readily estimate the Talairach coordinates using routines developed by Matthew Brett [113].

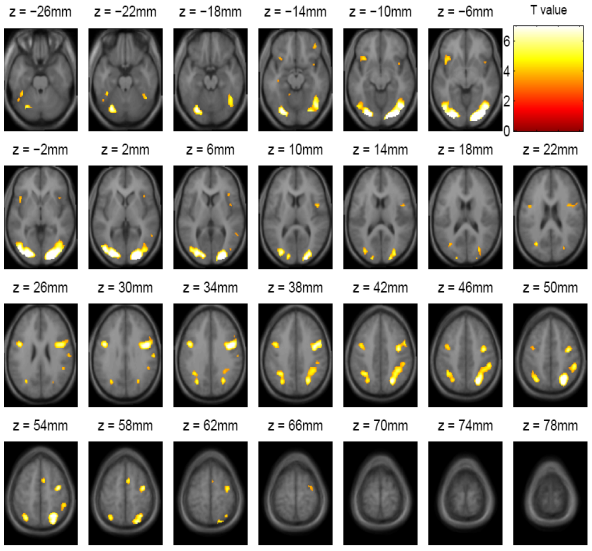
Table 4.1 and figure 4.1 show the activations found during the letter searching tasks. The activations were found in bilateral extrastriate cortices, superior parietal lobes, inferior parietal lobes, fusiform gyrus, prefrontal frontal lobes, primary motor area, supplemental motor areas and the cerebellum.



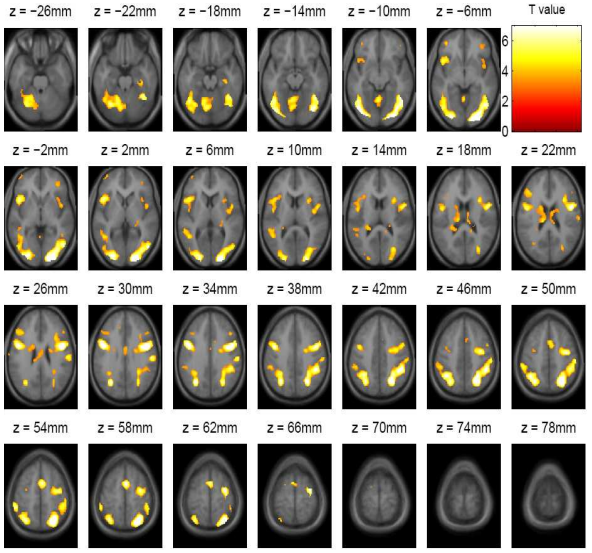

(a) HFW Task




(b) LFW Task




(c) LNW Task




(d) INW Task

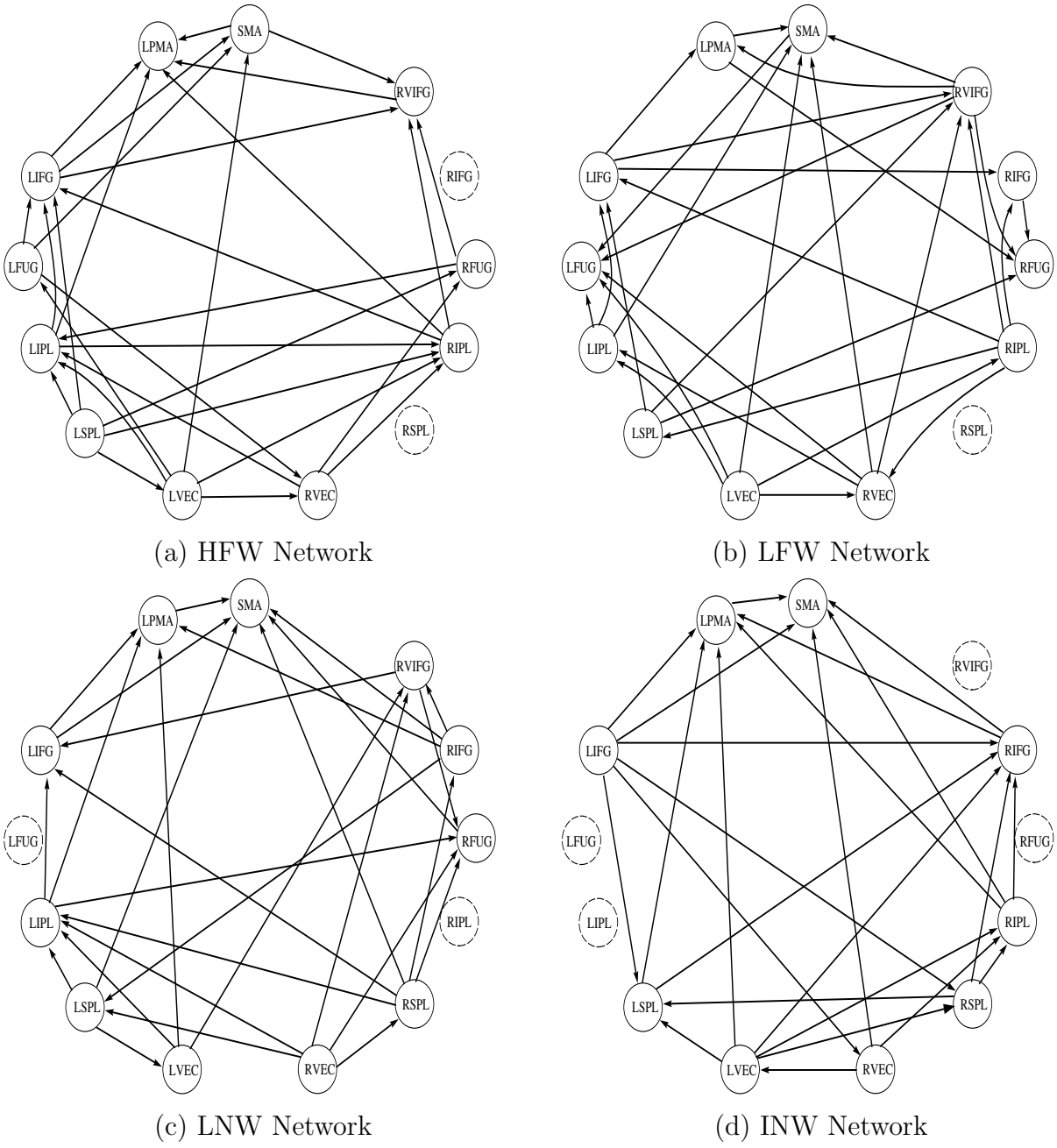
Figure 4.1: Brain regions showing significant activation in letter searching tasks relative to the rest condition: (a) the high frequency word reading, (b) low frequency word reading, (c) legal non-word reading and (d) illegal non-word reading. Statistical inferences were made at $p < 0.05$ corrected for multiple comparisons using FWER.

### 4.3.3 Derivation of Neural Systems

The time courses of significantly activated regions were extracted by taking the averages of the time-series at the peak-activated voxels and its neighbors (8mm) at the voxel-level for all subjects. The same procedures as described in the previous Chapter 3 were followed for driving the network structures from the extracted time-series. The networks which had the highest BIC score are shown in figure 4.2.

The extrastriate cortex (EC: BA18, BA19) in the visual cortex plays the role of visual representation in word processing [96]. The connection from the extrastriate cortex to superior parietal lobe (SPL: BA7) forms the dorsal stream of visual analysis, performing the perception of visual word form. As seen in figure 4.2 for most of network structures, the connections from EC to SPL are found in both hemispheres (LVEC → LSPL, RVEC → RSPL, LVEC → RSPL, and RVEC → LSPL). Meanwhile, the connections from the EC to prefrontal cortex including inferior frontal gyrus (IFG: BA44, BA45) and ventral inferior frontal gyrus (VIFG: BA47) represent the information flow for the processing of semantic analysis and decision (LEC→RVIFG, LEC→RIFG, REC→RVIFG, REC→LVIFG and REC→LIFG) [47]. Furthermore, the connections between EC and fusiform gyrus (FUG: BA37), associated with the retaining and recalling of words from the memory [96] are found in both hemisphere with reversed directions (REC→RFUG, LFUG→LEC); the reversed direction may be due to the bi-directional characteristic of the connectivity, represented by the Bayesian networks. In addition, a homologous interhemispheric connection between the ECs of both sides (REC→LEC) is found, which may be due to the transcallosal inferences between two hemispheres [51].

Parietal lobe generally performs the integration of sensory information for the




(a) HFW Network (b) LFW Network

(c) LNW Network (d) INW Network

Figure 4.2: Structures learned from the data for (a) the high frequency word reading, (b) low frequency word reading, (c) legal non-word reading and (d) illegal non-word reading. A dotted circle indicates that the region is not significantly activated in the particular task. L(R)VEC: left (right) ventral extrastriate cortex, L(R)SPL: left (right) superior parietal lobe, L(R)IPL: left (right) inferior parietal lobe, L(R)FUG: left (right) fusiform gyrus, L(R)IFG: left (right) inferior frontal gyrus, VIFG: ventral inferior frontal gyrus, PMA: primary motor area, SMA: supplemental motor area.

control of movement. In particular, the superior parietal lobe (SPL: BA7) plays the role of visual analysis and mainly makes efferent connections to the prefrontal cortex, providing more elaborate information (LSPL→LIFG, LSPL→RIFG) [96]. A homologous interhemispheric connection is also found between the SPLs (LSPL→RSPL). As seen in figure 4.2, the functional links from EC via SPL to prefrontal cortex form the dorsal visual pathway of language processing (LEC → LSPL → LIFG, LEC → LSPL → RIFG) [51]. The LIPL or supramarginal gyrus (BA40) activation has been linked to the memories of visual word forms in language system. In the derived networks, it is mainly connected with the occipital lobe and superior parietal lobe performing the initial processing of visual words.

The LIFG is mostly active for phonemic decisions and receives inputs from parietal lobes [97, 99]. In figure 4.2, the LIFG in most of networks has output connections to primary motor area (IFG → PMA). The VIFG (BA47), including orbitofrontal cortex, plays a specific role in controlling voluntary goal-directed behavior [26]. The connections of VIFG → LPMA (BA4) stores the voluntary activities involved [114, 115].

The SMA is believed to play a role in the planning of complex and coordinated movements [96]. The PMA is treated as the storage of motor patterns and voluntary activities and is involved in the expressive language of lips and tongue areas, and writing and sign language of hand and arm areas [114]. The connection LPMA → SMA is common for all tasks, indicating the voluntary movements involved in letter searching [115].

In the derived interhemispheric language network, the left hemisphere showed

dominant pathways, which is consistent with the traditional language network (left-hemispheric). Although activations have been symmetrically distributed in both hemispheres, the right hemisphere activations may be due to the transcallosal influence of the left.

The connections in the model that are consistent with previous literature are given in table 4.2. Due to the fact that the specific networks for each cognitive task are different even though the tasks are very similar (e.g., different presenting rate, different words, or different block design in reading tasks), existing literature can only be used as a general reference to an existing connection. The connectivity pattern derived from the present method is consistent with the information flow in the letter searching task as evidenced by the literature, but the connections without a corresponding reference cannot be refuted.

### 4.3.4 Derivation of Language System

Past neuroimaging studies have demonstrated that there are differences in the neuroanatomical representation of English at the levels of orthography, phonology and semantics. Evidence from these studies suggest that the differences lie in the location of the brain regions activated, i.e., orthographic processing is mainly subserved by the occipital regions, phonological processing is subserved by the left fronto-parietal regions and semantic processing by the left fronto-temporal regions.

In the experiments, fMRI data was collected in a letter searching task, which was designed to have a cross comparison among reading of high frequency word (HFW), low frequency word (LFW), legal non-word (LNW) and illegal non-word (INW), that

Table 4.2: List of the connections among activated brain regions, found to be involved in the reading task, which had been previously verified in other language-based tasks.

| Connection | Functional Discription | Relative Reference |
|---|---|---|
| LEC → LSPL | perception of visual word form | Horwitz [102] |
| REC → RSPL | perception of visual word form | McIntosh [51] |
| LEC → LIPL | memories of visual word forms | McKiernan [69] |
| LEC → LFUG | retaining and recalling of words from the memory | Kolb [96] |
| REC → RFUG | retaining and recalling of words from the memory | McIntosh [51] |
| LEC → REC | homologous interconnection | McIntosh [1] |
| | | McIntosh [51] |
| | | Krause [64] |
| LSPL → LFUG | semantic processing | Price [99] |
| | | Horwitz [102] |
| LSPL → RSPL | homologous interconnection | Honey [67] |
| LIPL → LIFG | executive control | McKiernan [69] |
| | | Petersson [70] |
| LFUG → LIFG | semantic phonologic retrieval | McKiernan [69] |
| | | Matsumoto [103] |
| LIFG → LPMA | motor function | Krause [64] |
| LIFG → SMA | motor function | He [116] |
| LIFG → RVIFG | inner speech production | Krause [64] |
| | | Nyberg [65] |
| | | Petersson [70] |
| LIFG → RIFG | homologous interconnection | Honey [67] |

will probably reveal the differences between the three components of language processing. Basically in the design, illegal non-word reading reveals the orthographic processing. When reading an illegal non-word, readers concentrate on the shape of characters and form of the words, hence the absence of phonological and semantic processing. High frequency word reading usually trigger semantic processing because these words are highly used and highly associated (context wise), and when reading them, there is little need to go through how it sounds (phonology processing) or what it looks like (orthography processing). Legal non-word and low frequency word reading would largely involve phonological processing, this being especially true for legal non-words as these words can be pronounced but have no meaning whatsoever (no semantics processing). Since low frequency words do not occur often, readers often need to pronounce them first (phonology processing) before attempting to process the words semantically.

As there will be a great deal of overlap in terms of brain activation for the three subcomponents of language processing, it is worthwhile to discuss the model to the three components collectively. This is the reason that the experiments were designed to have four tasks instead of three (one task for each subcomponent). The tasks lead to a gradual switch between the language processing components as it is not possible to truly isolate any one particular component in language processing. Thus low frequency word reading could also involve semantics processing and legal non-word reading might be effected by orthographic processing.

In this section, three neural systems referring to the language components are discussed individually, based on the models derived from the fMRI experiments. Further more, a possible language system that concludes the three subsystems is developed.

### Orthography System

The proliferation of neuroimaging studies on orthographic processing in English unilinguals makes a summary difficult [117]. Researchers have used a variety of tasks and found activations in many different brain regions. Experimental paradigms that have been developed to tap orthographic processing include single word reading [118, 119, 120, 121, 122], verbal fluency or word generation [123, 124, 125, 126], and case judgement on letter strings [127, 128]. Joseph *et al.* [117] noted that many of the tasks that have presumed to tap orthographic processing might also involve phonological decoding and even automatic activation of semantic processing. For example, in the word generation task, participants are presented with a single letter and instructed to generate a word that begins with the given letter. Although access to the orthographic lexicon is required in order to generate a word, participants may use phonological strategies by converting the given visual letter into a corresponding sound before retrieving words from the phonological lexicon (see Friedman et al., 1998). Also, other neuroimaging studies have suggested that the left frontal gyri (BAs 44, 45) contribute to the semantic rather than orthographic processing of words [129, 130, 131, 132]. Thus, Joseph et al. [117] concluded that there might be a great deal of overlap in terms of brain activation for both orthographic and phonological processing even though the tasks sought to tap only orthographic processing. Non-word reading networks in letter searching task were used to derive the orthographic system; particularly, illegal non-word reading network was mainly used with the assistance of legal non-word reading network by considering the overlapping problem mentioned above.

Figure 4.3 shows the proposed neuronal network for the orthographic system based

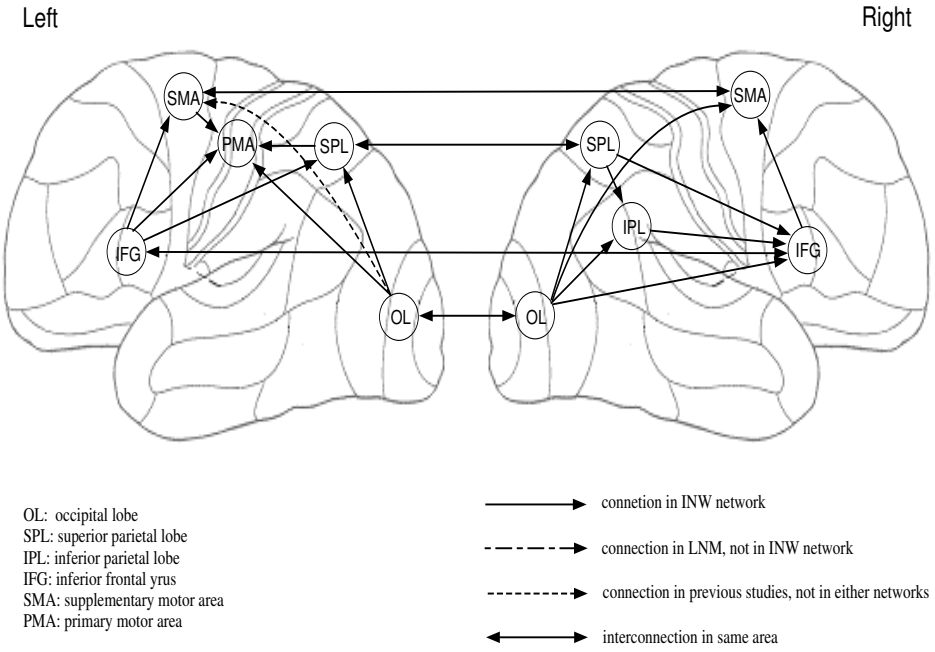



Figure 4.3: Schematic diagram showing the proposed neuronal network for the orthographic processing of English words during reading.

on the derived illegal non-word network (INW network) with the reference of legal non-work network (LNW network) and previous studies. The homologous regions including occipital lobes, parietal lobes, inferior frontal lobes, supplemental motor areas are interconnected. The initial stage of word reading involves early visual processing, which takes place bilaterally in the occipital lobe including the striate (BA 17) and extrastriate cortices (BAs 18 and 19). The information is sent to the superior parietal lobe (BA 7) for the perception of visual word form. In addition, the words viewed may automatically activate networks related to semantic processing and this is subserved by brain regions located in the inferior frontal lobes (BA 44/45). Finally, it is also likely that the reader would engage in phonological monitoring or

'subvocal articulation' during reading. This involves the premotor regions (i.e., medial BA 6 or supplementary motor area) and the motor cortex (BA 4).

**Phonology System**

The activation areas identified for phonological processing of English words vary in previous research. Again, the discrepancies in activation foci may be due to the diverse range of experimental paradigms employed by investigators. Access to the phonological components of language has been studied with tasks that require the perception and evaluation of the sound structure of words and letters [117]. These have included rhyme judgements [117, 120, 127, 133], passive word listening [134, 135, 136], phonological monitoring [137, 138], nonword reading [139, 140], etc. Legal non-word reading network in letter searching task with the assistance of low frequency word reading network, was used to derive the phonology system.

Figure 4.4 shows the proposed neuronal network for the phonology system based on the legal non-word network (LNW network) with the reference of low frequency word network (LFW network) and previous studies. In addition to the connections related to visual and orthographic processing (as discussed in the orthographic system), phonological processing also involves the inferior parietal lobe (BA 39,40) which has been linked to the memories of visual word forms in language system as the reader is trying to distinguish the legal non-word with formal word. Thus the information from occipital lobe and superior parietal lobe is sent to the inferior parietal lobe. The temporal lobe (BA 21,22,37) is also included, which likely involves the conversion of graphemes to phonemes. Finally, it is also likely that the reader would engage in phonological monitoring or 'subvocal articulation' during reading. This involves the

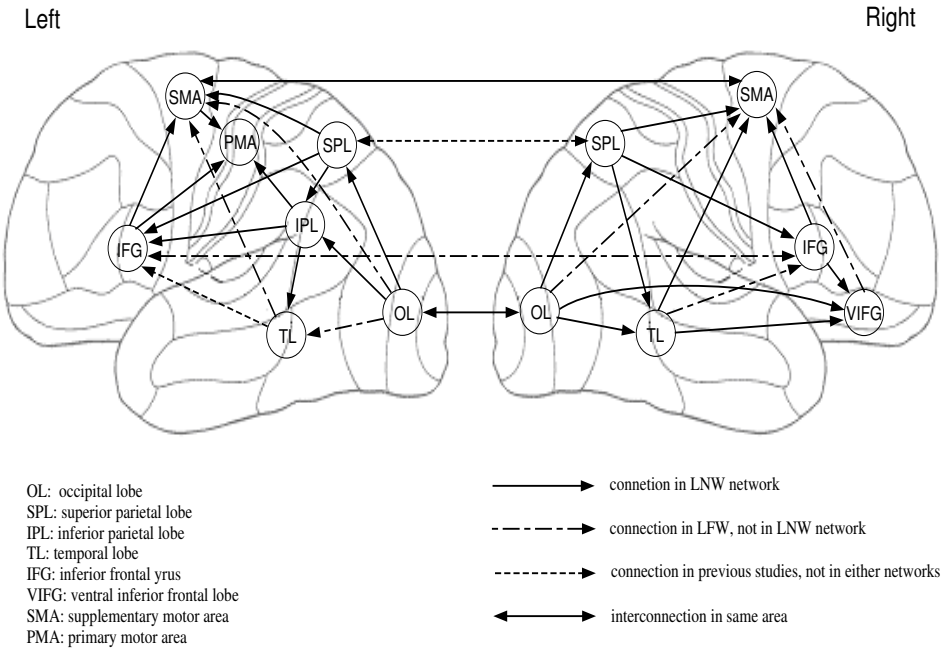



Figure 4.4: Schematic diagram showing the proposed neuronal network for the phonology processing of English words during reading.

premotor regions (i.e., medial BA 6 or supplementary motor area) and the motor cortex (BA 4).

### Semantics System

Experimental paradigms that have been developed to isolate the semantic processing of English words usually require participants to make judgements concerning word meaning or determine if the stimulus is a member of a designated category [117]. These areas were identified using a variety of experimental paradigms that included word generation [119, 120, 141, 142], categorical judgements [127, 137, 143], semantic monitoring, such as verb-noun comparison [105, 131, 136, 144], etc. Recent studies

also suggest that the posterior temporal regions may be responsible for the maintenance of semantic knowledge, whilst the inferior prefrontal regions may be involved in the executive processes of retrieving semantic information [145]. High frequency word reading network in letter searching task with the assistance of low frequency word reading network was used to derive the semantics system.

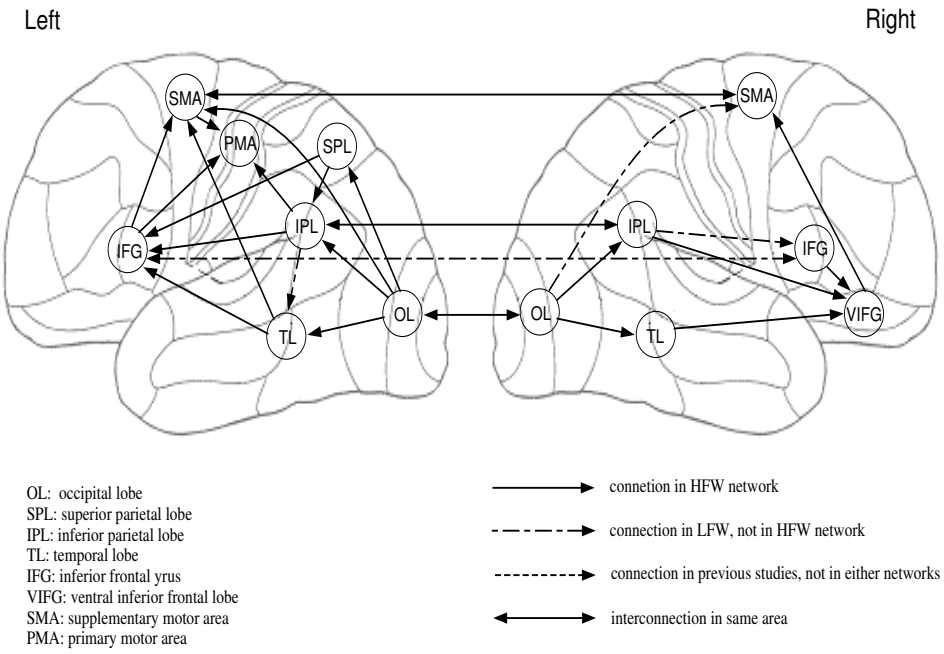



Figure 4.5: Schematic diagram showing the proposed neuronal network for the semantics processing of English words during reading.

Figure 4.5 shows the proposed neuronal network for the semantics system based on the high frequency word network (HFW network) with the reference of low frequency word network (LFW network) and previous studies. Thus, the neural network for the semantic processing of words is likely to connect the regions involving visual and orthographic processing (subserved by the regions in the ventral extrastriate cortices),

regions involving initial processing of visual word form (subserved by the parietal lobes) and the regions involving semantic monitoring and retrieval (subserved by the inferior prefrontal and middle temporal regions).

**Langauge System**

The process of reading generally involves a) visual processing, b) orthographic processing, c) phonological recoding, d) semantic retrieval and e) execution of motor plans. When a word is presented, visual processing takes place bilaterally in the striate (BA 17) and extrastriate cortices (BAs 18 and 19). In these areas, the fine-grained analysis and identification of the letter shapes take place, i.e., helping to distinguish if the letters are in upper or lower case, whether the letters are real alphabetic letters or meaningless shapes/symbols. After the information regarding the letter form has been established, orthographic processing takes place for perception of visual word form in superior parietal lobe (BA 7). After which, phonological processing may occur in the angular gyrus (BA 39). After the reader has established both the orthographical and phonological cues from the letter strings, the next step would be to perform a lexical search for the word's meaning in the mental lexicon. This is likely to involve the anterior frontal lobes (BA 44/45 or Broca's area), together with BA 37 (inferior posterior temporal lobe) and BA 40 (supramarginal gyrus in the inferior parietal lobe). Finally, it is also likely that the reader would engage in phonological monitoring or 'subvocal articulation' during reading. This involves the premotor regions (i.e., medial BA 6 or supplementary motor area) and the motor cortex.

Figure 4.6 concludes the schematic diagram showing the proposed neural network model for orthographic, phonological and semantic processing of English words during

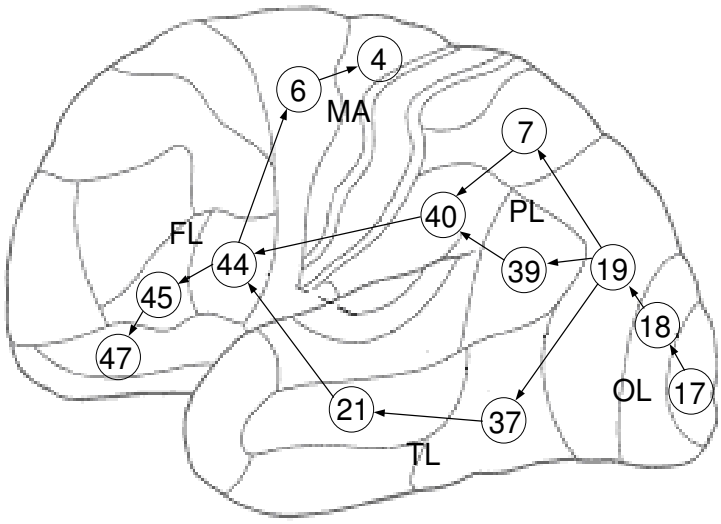



Figure 4.6: Schematic diagram showing the proposed neuronal network for orthographic, phonological and semantic processing of English words during reading. FL = frontal lobe; MA = motor area; TL = temporal lobe; PL = parietal lobe; OL = occipital lobe.

reading.

## 4.4 Discussion

Given that reading is a complex cognitive process which involves more than reading single words, it may involve a more extensive network than what has been described. However, as newer computational analyses and neuroimaging techniques such as Diffusion Tensor Imaging (DTI) emerge, we will be able to verify and validate if the proposed neural network model of reading accurately depicts reading in the healthy human brain. More importantly, we will then be able to conclusively determine the direction of communication between the identified brain regions.

Many tasks were designed to examine the access and the functioning of the language processing, yet the specific brain regions related to the processing of the three language components vary considerably. This disparity in the findings might be attributable to differences in experimental design, stimuli (words, pseudowords, letter strings, etc.), baseline tasks (fixation, line orientation judgement, etc.), and imaging method (PET, fMRI, etc.) across studies. In this experiment, fMRI data was collected from letter searching tasks. It is not possible to cover all the aspects mentioned above by a single experiment. More experiments should be used to cover other aspects or how these three processes cooperate and compete.

The complexity of the brain makes it difficult to be explored, especially in higher cognitive tasks; the analysis of functional integration (functional connectivity and effective connectivity) is still far from settled. The proposed method of exploring global neural systems from functional imaging data provides an alternate method to study brain function in terms of networks. The derived network is consistent to most of the previously published results.

# Chapter 5

# Modeling Brain Disconnectivity

The present method, which learns the structure of the functional network of the brain from functional MR images, has been illustrated in previous chapters. This chapter extends the method for applications in lesion studies. Firstly, how the networks derived from patient and healthy participants could be compared to explore disconnectivities of brain disease is illustrated. The approach is demonstrated by applying the method on fMRI datasets collected in a counting Stroop task. Secondly, the present method is extended by estimating the strengths of the interactions to explore the behavior of such networks due to an abnormal event. The latter is demonstrated by using a case study of stroke patient.

## 5.1 Introduction

This chapter consists of two parts, each showing a possible way of handling brain lesions by using Bayesian network. The first part shows how the networks derived from patient and healthy participants could be used to explore disconnectivities of brain disease. The second part shows how the behavior of neural networks due to an abnormal event is explored by estimating the strengths of connectivity.

Many researches have been done for the lesion study based on functional imaging techniques. Previous methods of disconnectivity detection are mostly comparing the activation patterns of patients and healthy controls. However, there are diseases that are due to disconnections, for example: the functional disconnection of the left angular gyrus, which is involved in normal reading task, has been found in the disease of dyslexia [21, 22]; the functional disconnectivity of the medial temporal lobe in Asperger's syndrome [23]; the functional disconnectivity in subjects at high genetic risk of schizophrenia [24].

The new method can be used in the analysis of disorders by studying the difference of the networks explored from fMRI data between patients and healthy participants performing the same cognitive tasks. The primary idea is that since Bayesian networks are able to derive the brain structures from functional data, the results from a patient and a normal person doing the same tasks should show the differences in the connectivity pattern. The method is demonstrated by exploring the functional structure from fMRI data obtained in a counting Stroop task.

The strength of the interactions between regions is another important factor for connectivity analysis. A brain disorder could be due to either a disconnectivity or just a weak connection. Taking this into consideration, the method can be extended to brain lesion study. The present method uses conditional probability densities (CPD) that describe the behavior of a network in a complete statistical sense, whereas the conditional probability that reveals how much a region's activation depends on the others represents the strength of an interaction. The approach is demonstrated in a case study of a stroke patient.

## 5.2 Disconnectivity analysis - X-syndrome patients with Stroop task

The present approach of global disconnectivity analysis benefits from the complete statistical representation of activity. The method is demonstrated by exploring the functional structure from fMRI data obtained in a counting Stroop task. In the experiment, the neural systems were derived for neutral and interference counting Stroop tasks performed by both normal control subjects and patients. The networks were used to infer the differences of the performances between groups. The technique is illustrate with an experiment on a fMRI dataset obtained from the fMRI Data Center [85], a counting Stroop task (access number: 2-2000-1123B). The structures of the neural systems involved in the tasks were derived and compared between the two groups.

### 5.2.1 Experiments and Results

#### Data

Functional MRI data used in this experiment was obtained from a counting Stroop task testing the cognitive interference that occurs when processing of one stimulus feature impedes the simultaneous processing of a second stimulus attribute [146]. The data was collected by Tamm *et al*[26] to investigate the performance of females with fragile X-syndrome on the cognitive interference task compared to a healthy control group. The participants included 14 females with fragile X-syndrome and 14 age-matched healthy control females without the fragile X mutation, ranging in age from 10 to 22 (mean age 15.43). The task consisted of 12 alternating experimental

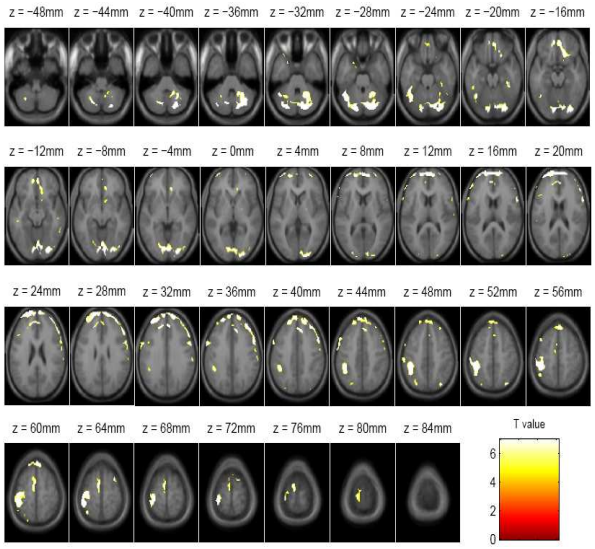
(interference) and controlled (neutral) conditions with the rest condition. For both conditions, subjects were instructed to press the button that corresponded to the number of words appeared on the screen. During the neutral counting task, the word "fish"was presented 1, 2, 3, or 4 times on the screen (15 trials) and during the interference counting task, the words "one", "two", "three"and "four"were presented 1, 2, 3, or 4 times on the screen (15 trials). Stimuli were presented for 1350ms at a rate of one every 2s (TR) for a total of 180 trials (90 experimental, 90 experimental, 90 control). For more details on the experiment, the reader is referred to Tamm *et al* [26].

### Detection of Activation

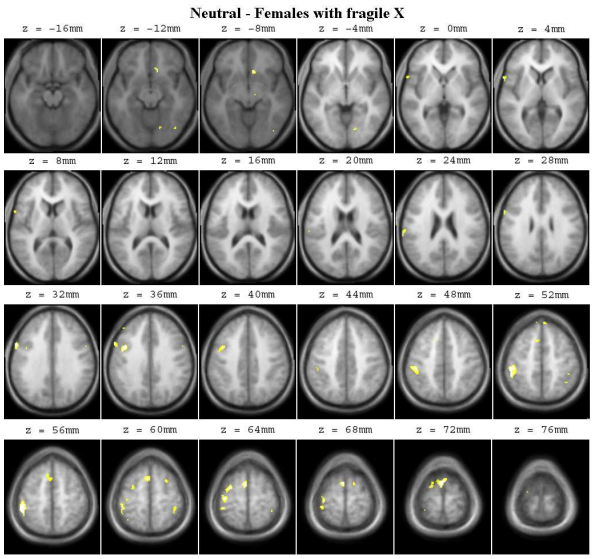
The preprocessed functional images of the subjects had been provided by the fMRIDC; images were reconstructed by using Inverse Fourier Transform for each of the 225 time points into $64 \times 64 \times 18$ image matrices and voxel size of $3.75 \times 3.75 \times 7$ mm. The regions showing increased activity during counting for both neutral and interference were identified by looking at the fMRI signal while counting relative to the rest condition. The changes in the blood oxygenation level dependent (BOLD) contrast, associated with the performance of the counting task, were assessed on a voxel-by-voxel basis by using the general linear model and the theory of Gaussian fields. The method takes the advantages of multivariate regression analysis and corrects for temporal and spatial autocorrelations of the fMRI data. Group analyses were performed using the fixed-effect analysis (FFX) which infer "typical"characteristics about the group of subjects.

Significant hemodynamic changes for each contrast were assessed using the $t$-statistical parametric maps and the results were reported by giving the $t$-values; And the statistical inferences were made at $p < 0.05$ corrected for multiple comparisons by using Family-Wise Error Rate (FWER). Both females with fragile X-syndrome and control groups did the interference counting and neutral counting tasks, there will be four groups of results ready for network analysis: interference-control (IC, control group doing interference task), neutral-control (NC, control group doing neutral task), interference-fragile X (IFX, females with fragile X-syndrome group doing interference task) and neutral-fragile X (NFX, females with fragile X-syndrome group doing neutral task). The aim of this experiment is to explore the functional networks from data for different groups performing the interference counting and neutral counting tasks and make inferences about the differences.

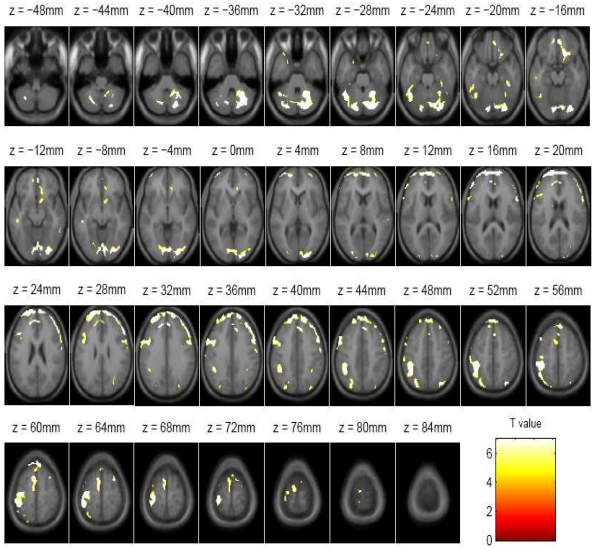
SPM2 was used for the above analysis. Table 5.1 and figure 5.1 show significant activation for each group in this experiment. Most of the significant activations had been found in bilateral parietal and frontal lobe, including superior parietal lobe (SPL), inferior parietal lobe (IPL), superior frontal gyrus (SFG), middle frontal gyrus (MFG), inferior frontal gyrus (IFG), primary motor area (PMA), supplementary motor area (SMA) and anterior cingulate cortex (ACC). A comparison between the two groups (females with fragile X-syndrome and control group) reveals that there are more activations in inferior parietal lobe (BA40) and motor areas (SMA BA6) for the females with fragile X-syndrome group, and much stronger activity in middle frontal gyrus including Brodmann Area 9 and 10, which are treated as reasoning judgement areas, in the control group. Although proper motion correction was performed on the data, the crescentic frontal activations (AFG) in figure 5.1 may looked like motion



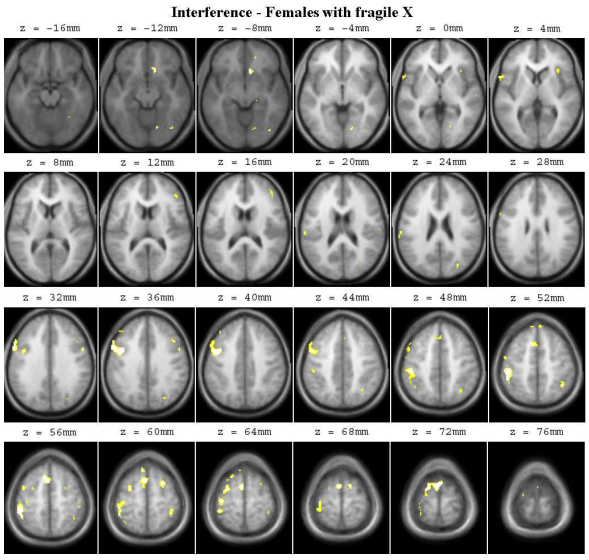

(a) NC Task




(b) NFX Task




(c) IC Task




(d) IFX Task

Figure 5.1: Brain regions showing significant activation in all counting tasks relative to the rest condition: (a) the neutral-control task, (b) the neutral-fragile X task, (c) the interference-control task and (d) the interference-fragile X task. Statistical inferences were made at $p < 0.05$ corrected for multiple comparisons using FWER.

Table 5.1: The results of the analysis of brain activation patterns. Significantly activated regions during the counting tasks relative to the rest condition are shown in 3D MNI coordinates with the significance given by $t$-values. Statistical inferences were made at $p < 0.05$ corrected for multiple comparisons by using FWER for NC (neutral-control), IC (interference-control), NFX (neutral-fragile X), and IFX (interference-fragile X) groups. Statistical inferences were made at $p < 0.05$ corrected for multiple comparison using FDR. (SPL: superior parietal lobe; SG: supramarginal gyrus (inferior parietal lobe); SFG: superior frontal gyrus; MFG: middle frontal gyrus; IFG: inferior frontal gyrus; PMA: primary motor area; SMA: supplementary motor area; ACC: anterior cingulate cortex; the letters L and R before the regions indicating the left and right hemisphere).

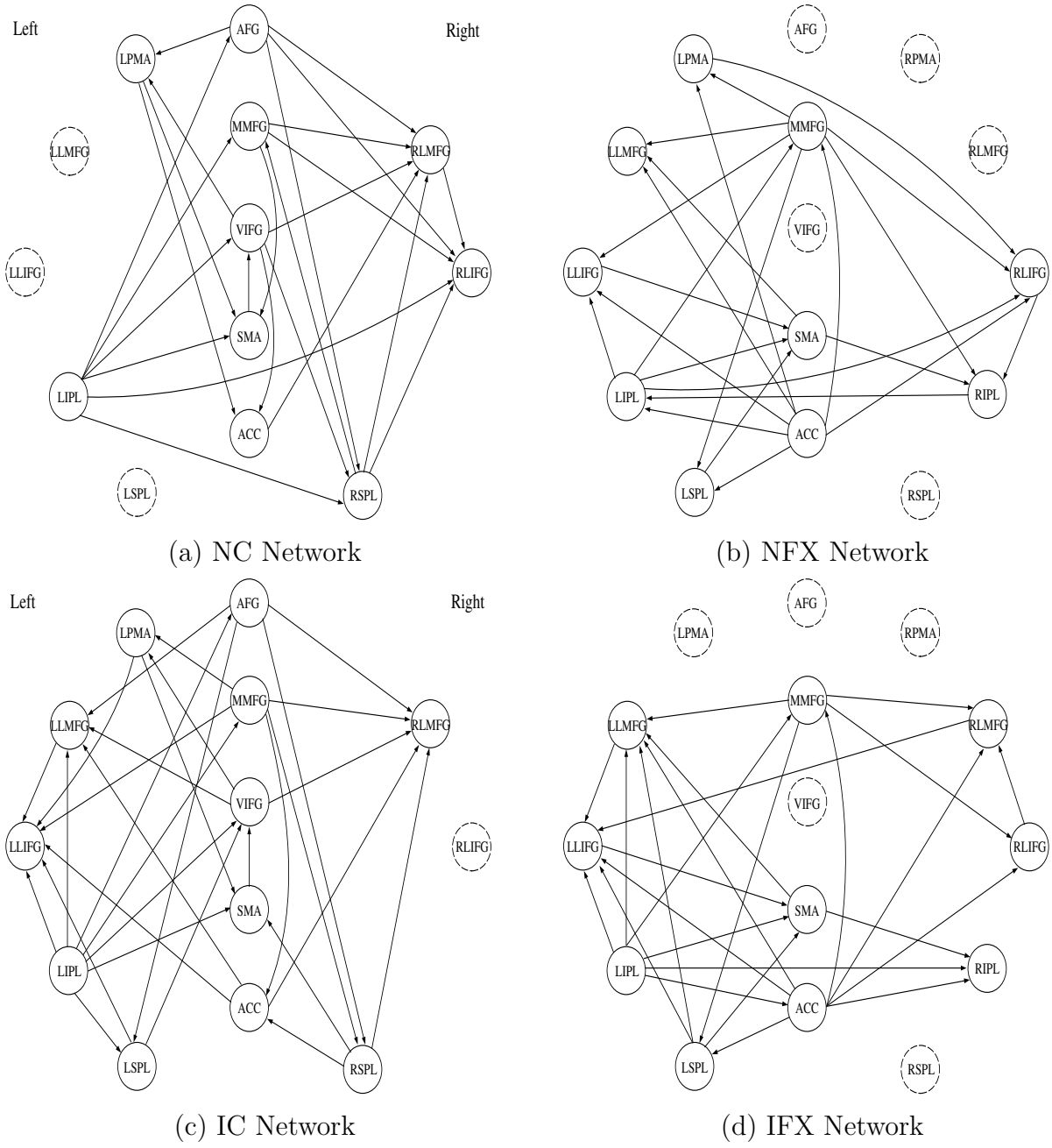
| Region | NC | | IC | | NFX | | IFX | |
|---|---|---|---|---|---|---|---|---|
| | Coordinates | $t$ | Coordinates | $t$ | Coordinates | $t$ | Coordinates | $t$ |
| LSPL(BA 7) | | | (-28,-74,50) | 9.99 | (-36,-56,60) | 8.65 | (-36,-56,60) | 7.53 |
| RSPL(BA 7) | (32,-72,50) | 7.34 | (32,-72,50) | 6.64 | | | | |
| LIPL(BA 40) | (-42,-38,58) | 7.02 | (-42,-38,60) | 8.63 | (-50,-36,54) | 6.35 | (-46,-46,58) | 9.45 |
| RIPL(BA 40) | | | | | (40,-40,62) | 7.73 | (40,-40,62) | 5.71 |
| LLMFG(BA 9) | (-16,38,32) | 8.66 | (-16,38,32) | 8.32 | (-10,34,62) | 9.95 | (-8,32,62) | 7.77 |
| RLMFG(BA 9) | (6,38,60) | 6.37 | (6,38,60) | 7.17 | (6,48,52) | 6.98 | (6,46,52) | 8.47 |
| MMFG(BA 8) | (-18,46,42) | 8.07 | (-16,48,42) | 6.91 | (-42,38,36) | 11.0 | (-40,36,40) | 8.94 |
| LLIFG(BA 44) | | | (-56,8,34) | 7.03 | (-60,12,32) | 8.32 | | |
| RLIFG(BA 44) | (56,8,34) | 6.86 | (58,14,18) | 9.76 | (56,8,34) | 5.79 | (56,8,34) | 7.03 |
| SMA(BA 6) | (-6,-4,64) | 9.27 | (-6,-4,66) | 7.78 | (-2,4,72) | 8.23 | (-2,4,72) | 5.91 |
| LPMA(BA 4) | (-32,-26,68) | 5.93 | (-34,-24,66) | 6.94 | (-38,-20,64) | 6.79 | (-38,-20,64) | 8.34 |
| ACC(BA 24) | (10,36,-8) | 9.03 | (10,34,-8) | 9.31 | (10,22,-10) | 7.05 | (10,22,-10) | 9.63 |

artifact [48, 147, 37, 148]. This region was still included in the model, as under proper preprocessing [37] and statistic analysis the region was significantly activated.

### Derivation of Neural System

The same procedures as described in the previous Chapter 3 were followed for driving the network structures from the extracted time-series. The networks which have the highest BIC score in each task are shown in figure 5.2.

The study regarding this experiment data conducted by Tamm *et al* [26] investigated the performance of females with fragile X on the cognitive interference task




(a) NC Network (b) NFX Network

(c) IC Network (d) IFX Network

Figure 5.2: Structures learned from data, each node represents a brain region and the arcs showing their connections, solid circle indicating that the region is not significantly activated in that particular experiment: (a) the neutral-control network, (b) the neutral-fragile X netwrok, (c) the interference-control network, and (d) the interference-fragile X network.

comparing with a healthy control group using fMRI. Females with fragile X syndrome is the most common form of inherited developmental and learning problems. They are known to be impaired in executive function. The study revealed that females with fragile X performed less well than controls in terms of both speed and accuracy on the counting Stroop task. Compared to healthy controls, females with fragile X show different patterns of activation, particularly in the prefrontal cortex, and a specific deficit in the orbitofrontal gyrus, as well as strikingly different patterns of deactivation, suggesting that females with fragile X have anomalous brain activation during cognitive interference processing tasks and may fail to appropriately recruit and modulate lateral prefrontal cortex and parietal resources [26].

**Network Study within a Group**

This section studies the networks between neutral and interference conditions for each group (control group and females with fragile X syndrome group). The similar activation seen in the medial cortices for the two conditions may indicate that the function of counting is mainly processed by the medial areas especially in the anterior cingulate cortex (ACC: BA24), which had been shown to be playing an essential role in counting Stroop [146, 149, 150]. So, the different activation in the lateral cortices between the neutral and interference conditions may reflect the effects of the "interference"; more activation in the language areas in the left hemisphere was found in the interference counting task. This is due to the fact that the subjects had been more attracted or interferred by the meaning of the words, being counted in the interference counting task.

The ACC (BA24) is engaged during the Stroop task in order to resolve competing streams of information by selection sensory inputs and responses [146]. The effects

are reflected in the IC network by the connections, ACC → LLMFG (BA9) and ACC → RLMFG (BA9), that involved in resolving interference effects, and ACC → LLIFG (BA44) for phonemic decisions. The difference of connection shows the more involvement of the semantic processing and decision making in the IC network.

The LMFG (BA9) is involved in the tasks that require executive control and receives input from posterior parietal, and the selection of behavior based on the short term memory [97, 99]. Particularly in this experiment, this region is involved in processing Stroop-related conflit and resolving interference effects [26]. The LLMFG in the IC network is connected to the LLIFG (BA44) and the RLMFG in the NC network has been connected to the RLIFG for executive controls. The MMFG (BA8) is believed to play an important role in the control of eye movements [114]. The common connections found for both tasks are MMFG → RLMFG, MMFG → RLIFG, and MMFG → LLIFG. The connection MMFG → SMA in the NC network is absent in the IC network while the connections, MMFG → RSPL, MMFG → ACC, and MMFG → LPMA, in the IC network are absent in the NC network. The difference may be due to the different concentration demanded by the tasks.

The LIFG (BA44) is mostly active for phonemic decisions and receives inputs from parietal lobes [97, 99]. In figure 5.2, the LIFG in both networks has no output connection to other region. The VIFG (BA47) including orbitofrontal cortex plays a specific role in controlling voluntary goal-directed behavior [26]. The common connection for both task, VIFG → LPMA (BA4), stores the voluntary activities involved in the task [114, 115]. The connection VIFG → LLMFG (BA9) that is found only in IC network, is related to the specific function of LMFG that involved in processing Stroop-related conflit.

The AFG (BA10) is believed to play a part in strategic processes involved in memory retrieval and executive function [114]. The connections from AFG to other regions that are common in both networks include: AFG → SPL (BA7) and AFG → LMFG (BA9). The connections, AFG → LPMA (BA4) and AFG → RLIFG (BA44) present in the NC network are absent in the IC network. The SMA is believed to play a role in the planning of complex and coordinated movements [96]. The connection is found from SMA to ventral inferior frontal gyrus in both networks. The PMA is treated as the storage of motor patterns and voluntary activities and is involved in the expressive language of lips and tongue areas, and the writing and sign language of hand and arm areas [114]. The connection LPMA → SMA is common for both tasks indicating the automatic movements involved in counting task [115].

The parietal lobe generally performs the function of processing and discriminating of the sensory inputs [96]. The LIPL or supramarginal gyrus (BA40) activation, which has been linked to the memories of visual word forms in language system, observed in this experiment likely resulted from the arithmetic computing (counting) and language processing (reading) involved in the task. As seen in figure 5.2, the LIPLs in both networks send the representations of the inputs to the medial regions, AFG, MMFG, VIFG, and SMA, which are mainly involved in the counting function. The differences are seen as the extra activations in the language areas of the IC network: the connections from LIPL to LLIFG (BA44) and LSPL (BA7); as well as the connection for processing Stroop-related conflit and resolving interference effects: LIPL to LLMFG (BA9). The connections from LIPL to RSPL (BA7) and RLIFG (BA44) are seen only in the NC network; this may account for a compensation function for the absent of language pathways that present in the IC network, and be more likely

involved in visualization of symbols instead of reading, i.e., the ″automatic speech″, where the right hemisphere is subserving residual aphasia speech [151]. The LSPL (BA7) was found to be activated only in the interference counting task and has connections to the regions, LLIFG and VIFG (BA47); the RSPL was activated in both tasks and connected to the RLMFG (BA9) while the connections from RSPL to SMA (BA6) and ACC (BA24) are only found in the IC network.

In summary, the structures involved in both tasks are mostly common and the differences are mainly due to the specific language areas activated in the interference counting task. The connections, such as LIPL → LLMFG → LLIFG, which are only present in the IC network construct the language pathway thus performing the phonetic and semantic analysis and decision. Meanwhile, the connections, such as LIPL → RLIFG, which are only found in the NC network may perform a compensational function for the non-activated patterns corresponding the connections, LIPL → LLIFG and LIPL → SPL present in the IC network. In addition, based on the fact that the interactions between two regions could be bi-directional, some connections are seen reversed between the two networks such as MMFG → RSPL in the IC network versus RSPL → MMFG in the NC network.

The networks do not show obviously different between neutral and interference conditions for the females with fragile X syndrome group. The effects of ″interference″ are not seen from the networks, indicating the abnormal cognitive interference processing for the female with fragile X.

**Network Study between Groups**

This section describes the differences in networks derived from the two groups (patient group and control group) for each of the conditions (interference condition

and neutral condition).

Females with fragile X syndrome is the most common form of inherited developmental and learning problems. They are known to be impaired in executive function. The performance of the experiment revealed that females with fragile X performed less well than controls in terms of both speed and accuracy on the counting Stroop task. Compared to healthy controls, females with fragile X show different patterns of activation, particularly in the prefrontal cortex including the anterior frontal and the primary motor area, and a specific deficit in the orbitofrontal gyrus.

As mentioned in the previous section, the ACC (BA 24) played an essential role in counting task. The region in both network has the connections with SPL (BA 7), LIFG (BA 44) and MMFG (BA 8), however, the key connection ACC → LLMFG (BA9) which reveals the resolving of interference effects is not presented in the IFX network. LMFG (BA 9) is another important region that involved with executive control. Although the region in IFX network still receives input from parietal lobe, the connections for processing the executive control: LLMFG → LLIFG (BA 44), AFG (BA 10) → LLMFG and VIFG (BA 47) → LLMFG which are presented in the IC network, are missing in the IFX network. The LLIFG (BA 44) does not receive input from parietal lobe as the connections presented in the IC network. The AFG (BA 10) and VIFG (BA 47) are not activated in the IFX network, and hence the relative connections in the IC network are not presented in the IFX network.

Although at the first glance the networks between two groups have similar connections, some of the important connections that reveal the interference effects in the IC network are not presented in the IFX network. This fact reflects the conclusion made

by the previous literature [26], that females with fragile X have anomalous brain activation during cognitive interference processing tasks and may fail to appropriately recruit and modulate lateral prefrontal cortex and parietal resources.

The networks between groups for neutral task are less in common due to the fact that the NFX network is actually similar to the IFX network, indicating that female with fragile X-syndrome has abnormal effects between neutral and interference task. The interference processes had been actually merged between the neutral and interference tasks. As a result, the NFX and IFX networks are likely "sitting"in the middle between NC and IC networks.

### 5.2.2 Discussion

The complexity of the brain makes it hard to be explored especially in higher cognitive tasks; the analysis of the functional integration (functional connectivity and effective connectivity) is still far from converging. The proposed new method of exploring global neural systems from functional imaging data provides an alternate method to study brain function in terms of networks. The inference network derived showed the involvement of language areas in the interference counting task.

As illustrated in the experiment, the present method offers the feasibility of comparing differences of the networks performing different tasks. This could be used to differentiate the performance of patients and healthy participants performing the same cognitive tasks, and hopefully, provide a way to explore the disconnectivity hypothesis in brain disease.

## 5.3 Estimating the Strength of Connectivity - case study of stroke patient

In this experiment, the neural structure of a stroke patient performing English homophone matching task was derived. The structure was compared with the network derived from a healthy control group doing the same task to explain the possible disconnectivity in neural system of the stroke patient. To this end, the connection strengths of connectivity between normal group and the stroke patient were compared.

### 5.3.1 Learning Parameters

For a set of brain regions $R$ and a given a dataset $D$, the likelihood of the parameters $\theta$ is given by

$$p(D|\theta) = \prod_{k=1}^{N} p(x^k|\theta) \tag{5.3.1}$$

The goal of learning is to find the values of the parameters of each CPD which maximizes the likelihood of the training data. Equivalently, we can maximize the log likelihood:

$$\begin{aligned} L(\theta) &= \sum_{k=1}^{N} \log p(x^k|\theta) \\ &= \sum_{k=1}^{N} \sum_{i=1}^{n} \log p(x_i^k|a_i^k, \theta_i) \end{aligned} \tag{5.3.2}$$

where the last equality makes use of the factorization (Eq. 3.2.3) of joint distribution in the directed graphical model. The log-likelihood scoring function decomposes according to the structure of the graph, and hence we can maximize the contribution to the log-likelihood of each node independently. Then the log-likelihood decouples into

a sum of local terms involving each node and its parents, $L(\theta) = \sum_{i=1}^{n} L_i(\theta_i)$ where

$$L_i(\theta_i) = \sum_{k=1}^{N} \log p(x_i^k | a_i^k, \theta_i) \tag{5.3.3}$$

Then, all that remains is how to estimate the local parameters, $\theta_i$, given its local data $\{(x_i^k, a_i^k) : k = 1, 2, \ldots N\}$.

Considering the multivariate Gaussian distribution for the neural system, i.e., $p(x) = N(\mathbf{m}, \Sigma^{-1})$ [152] [153], we have:

$$p(x) \quad = (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\} \tag{5.3.4}$$

where $\Sigma = \{\sigma_{ij}\}_{n \times n}$ is a covariance matrix, $\mu$ is the $n$-dimensional mean vector, and $|\Sigma|$ is the determinant of $\Sigma$. It is convenient to refer to the *precision matrix* $W = \Sigma^{-1}$, whose elements are denoted by $w_{ij}$. Then, the conditional distributions is written as a product of component distributions, each being an independent normal distribution:

$$p(x_i | x_1, x_2, \ldots, x_{i-1}) = N(\mu_i + \Sigma_{j=1}^{i-1} b_{ij}(x_j - m_j), \frac{1}{v_i}) \tag{5.3.5}$$

where $\mu_i$ is the unconditional mean of $x_i$, $v_i$ is the conditional variance of $x_i$ given values for of the set $\{x_1, x_2, \ldots, x_{i-1}\}$, and $b_{ij}$ is a linear coefficient reflecting the strength of relationship between $x_i$ and $x_j \in \{x_1, x_2, \ldots, x_{i-1}\}$. So, we can describe a multivariate normal distribution as a Bayesian network, where there is an arc from $x_j$ to $x_i$ whenever $b_{ij} \neq 0$ [153].

Given a multivariate normal density, we can generate a Gaussian belief network, and vice versa. The unconditional means $\mathbf{m}$ are the same in both representation. The general transformation from $\mathbf{v}$ and $b_{ij}$ of a given Bayesian network $S$ to the precision matrix $W$ of the normal distribution represented by $S$, was described by Shachter and Kenley [154]. They used the following recursive formula in which $W(i)$ denotes

the $i \times i$ upper left submatrix of $W$, $b_i$ denotes the column vector $(b_{1,i}, \ldots, b_{i-1,i})$, and $b'_i$ denotes the transposed vector $b_i$:

$$W(i+1) = \begin{pmatrix} W(i) + \frac{b_{i+1} b'_{i+1}}{v_{i+1}} & -\frac{b_{i+1}}{v_{i+1}} \\ -\frac{b'_{i+1}}{v_{i+1}} & \frac{1}{v_{i+1}} \end{pmatrix} \tag{5.3.6}$$

for $i > 0$, and $W(1) = \frac{1}{v_1}$. Hence, known structure $S$ and data $D$, the parameter $b$ can be computed.

### 5.3.2 Experiments and Results

#### Data

The data was obtained by Singapore General Hospital, Department of Magnostic Radiology. The patient and the normal person did the same language task called *Homophone Matching*. Imaging was performed on a 1.5 Tesla whole-body magnetic resonance imaging (MRI) scanner (Siemens Vision; Erlangen, Germany). Prior to imaging, participants were briefed on the scanning procedures and experimental conditions so as to minimize anxiety and enhance performance. Participants were asked to lie supine inside the MRI scanner with their heads inside a standard head coil. Head movement was minimized within the coil using foam wedges, and a restraining band was placed across the forehead. Participants were also fitted with headphones ($MSI^{TM}$, Tampa, Florida) that attenuated ambient scanner noise by 30dB, and provided a means of conveying the experimenter's instructions. The presentation of written words and characters was controlled by Eprime software running on an IBM-compatible computer located outside the scanning room. Using a block design, each run consisted of four periods of the same language task (i.e., activation condition

with homophone task), interleaved with four periods of fixation (i.e., baseline condition with fixation "+", presented for 30 seconds. All stimuli were back-projected via a high- resolution LCD projector ($MSI^{TM}$, Tampa, Florida) onto an opaque screen positioned at the head end of the bore, with a specially designed mirror mounted on the head coil for viewing. For the activation periods, participants were asked to respond to the words and the words and characters by pressing "yes"(right hand) if the stimuli sound alike or "no"(left hand) stimuli sound different. For the fixation (baseline) periods, participants were instructed to focus their entire attention on a fixation point ('+' sign) shown on the middle of the screen.

Prior to the fMRI scans, tri-planar scout images in the sagittal plane and T1-weighted 3D coronal anatomical images using the MPRAGE sequence were acquired for each participant. This procedure provided high-resolution images of the entire brain and subsequently served as the structural scans for Talairach transformation. Functional images were obtained with a T2-weighted gradient echo, echo planar imaging (EPI) sequence (slice thickness = 1 mm, gap size between the slices = 8 mm, TR/TE/$\theta$ = 3000ms/66ms/$90^o$, voxel size = $1.5 \times 1.5 \times 8.0$ mm), FOV = $240 \times 240$ mm and acquisition matrix = $128 \times 128$, using blood oxygen level-dependent (BOLD) contrast. Ten contiguous oblique, axial slices covering most of the brain were acquired.

### Detection of Activation

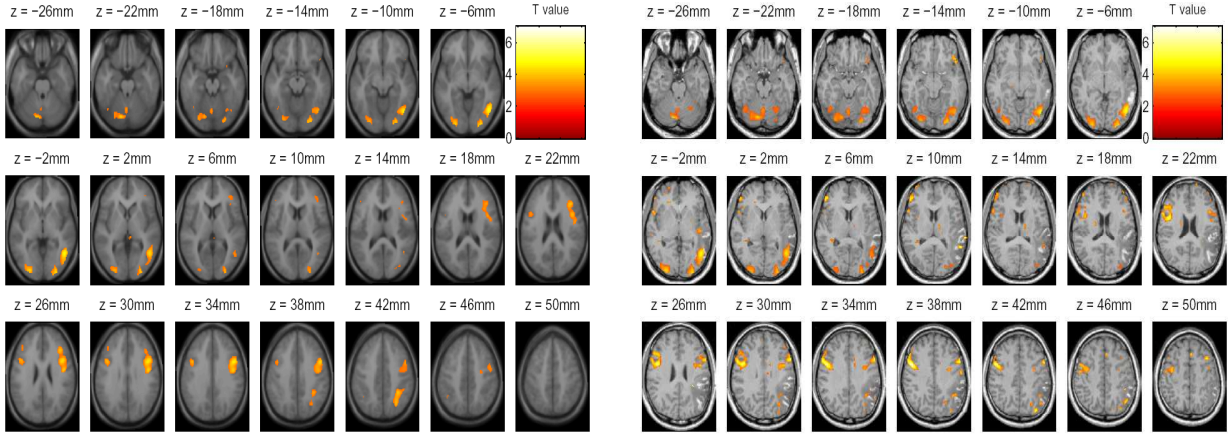
The fMRI data were analyzed using the SPM2. The first five scans (15 seconds) in each session (during which magnetization steady state was being reached) were excluded from the data analysis. All functional images were first corrected for movement using least-squares minimization, and then co-registered to the subjects 3D

T1-weighted image. Using the 3D image as a guide, the functional images were subsequently spatially normalized into the SPM standard space. Images were then re-sampled every 2mm using sinc interpolation, and smoothed with a FHWM 6mm, 3D Gaussian kernel to decrease spatial noise. Changes in blood oxygenation level dependent (BOLD) contrast associated with the performance of the reading tasks were assessed on a pixel-by-pixel basis, using the general linear model and the theory of Gaussian fields as implemented in SPM2. This method takes advantage of multivariate regression analysis and corrects for temporal and spatial autocorrelations in the fMRI data. Group analyses were investigated using fixed-effect analysis (FFX), which infers "typical"characteristics about the sample of participants. Each participants. Each language effect was tested by applying and significant hemodynamic changes for each contrast were assessed using the $t$-statistical parametric maps. Activations below a threshold of $p < 0.05$ corrected for multiple comparison using Family-Wise-Error-Rate (FWER) were reported.

Figure 5.3 shows the results for activation detection. Activations for task relative to rest were found for both normal group and patient in bilateral posterior fusiform, temporo-occipital, inferior and middle occipital, inferior parietal gyrus, left superior temporal gyrus, SMA, bilateral inferior frontal sulcus, bilateral middle frontal gyrus, bilateral superior frontal gyrus and cerebellum.

### Derivation of Neural System

Same procedures were used to extract the time course as the previous experiments. Despite deriving neural system using both left and right hemispheres previously, as the patient in this experiment has a stroke on the left hemisphere and according to




(a) Normal (b) Patient

Figure 5.3: Brain regions showing significant activation in English homophone matching tasks relative to the rest condition: (a) the normal group, (b) the patient. Statistical inferences were made at $p < 0.05$ corrected for multiple comparisons using FWER.

the traditional agreement that English language task is mainly involved with the left hemisphere, only the connectivity on the left hemisphere was considered to specify and simplify the disorder analysis. The regions of interest are: VEC: Ventral Extrastriate Cortex, (BA18,19) (-47 -60 -6); IPL: Inferior Parietal Lobule (BA39,40) (-24 -66 42); STG: Superior Temporal gyrus (BA22) (-43 -41 14); SMA: Supplementary Motor Area, (BA6) (-38 6 41); IFG: Inferior Frontal Gyrus, (BA44,45) (-47 27 -14).

The present method was implemented on both the patient and normal. The networks with the highest score for each participant were shown in figure 5.4 and figure 5.5.

Table.5.2 and Table.5.3 show the marginal probabilities of each node after training for each subject, where the conditional probabilities that represent the interactions have been estimated.

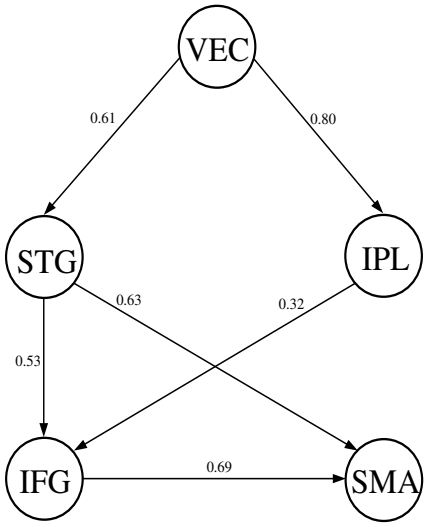



Figure 5.4: Neural system for the patient learned from a posteriori estimate.

### 5.3.3 Discussion

In this experiment, the patient has a wedge-shaped lesion in his right hemisphere. Theoretically, the left hemisphere is mainly involved in the English language task, and the Chinese language more occur in the right. The results of activation detection show: For the case that the patient doing English homophone matching (described in section 5.3.2), the activation in the left is basically normal except the weaker activation in parietal and temporal areas. Observed from the networks for both participants (figure 5.4 and figure 5.5), obviously, the connection between parietal and temporal is not found in the patient's structure. Besides, the connections in STG - IFG and IFG - SMA are stronger in the patient's structure, while connections in VEC - IPL and IPL - IFG are weaker. The results suggest that there are abnormal connections between temporal and parietal areas in this case study. The functional disconnectivity found correlated well with the patient's clinical symptoms of dysphasia.

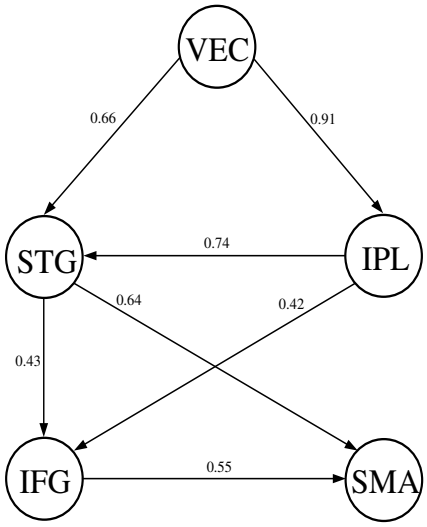



Figure 5.5: Neural system derived from the normal group.

## 5.4 Conclusion

Functional MRI is increasingly being introduced into clinical practice. Applying the present techniques into lesion studies will lead to a new way of looking into the abnormal brain. Brain disorders were analyzed by implementing the Bayesian networks individually on the fMRI scans from a patient and a normal subject group doing the same task. From the learning results, the difference of the structures and interconnections can be observed. The suggestions on the brain disorders can be made base on the observation.

In the second experiment, the patient has a wedge-shaped lesion in his right hemisphere. The activations do not show much difference between normal and patient participants, while results from the data-driven method suggest that there are connection problems with temporal and parietal areas in this case study. The functional disconnectivity correlated well with the patient's clinical symptoms of dysphasia. The

Table 5.2: Marginal probabilities for patient model: Marginal probabilities learned from patient's data. Each section contains the probability that a node is active given that the parent is active. Row list of variables are the current nodes (children). Column list are the parents. Cells containing a dash indicating that there is no connection in between.

| | VEC | IPL | STG | IFG | SMA |
|---|---|---|---|---|---|
| VEC | - | 0.80 | 0.61 | - | - |
| IPL | - | - | - | 0.32 | - |
| STG | - | - | - | 0.53 | 0.63 |
| IFG | - | - | - | - | 0.69 |
| SMA | - | - | - | - | - |

Table 5.3: Marginal probabilities for normal model: Marginal probability learned from normal model. Row list of variables are the current nodes(children). Column list are the parents. Cells containing a dash indicating that there is no connection in between.

| | VEC | IPL | STG | IFG | SMA |
|---|---|---|---|---|---|
| VEC | - | 0.91 | 0.66 | - | - |
| IPL | - | - | 0.74 | 0.42 | - |
| STG | - | - | - | 0.43 | 0.64 |
| IFG | - | - | - | - | 0.55 |
| SMA | - | - | - | - | - |

results again prove the hypothesis that there are diseases which is due to the connections. The disorder is not only caused from the disconnection but also the strength of connection.

Another major advantage of Bayesian networks might be its ability to infer network function in the case of brain disorders, as inferencing is a strength of the graphical models. By simulating the brain lesion study, by means of artificially abandoning some nodes (to mimic that some regions are not functional), or some connections (to simulate that some interactions between regions are disconnected), and then reconstruct the structure to simulate the brain abnormality. In a structure, each region

is responsible for a specific function or behavior. The abnormality at a structure as an event with a probability can be mimicked, so that the probabilities of activation of other structures can be predicted. The approach is not complete yet, we wish to propose this as a general frame work to analyze a wide range of brain disorders in the future research.

Exploratory connectivity is an attracted trend for network analysis. The used of graphical model is a start for exploratory analysis, nevertheless, it could be contributed to disconnectivity analysis for brain lesion. Brain disorders such as dyslexia are still not completely explored, the proposed technique aims for simulating and analyzing lesion studies by exploratory extracting the abnormal structure from fMRI scanning, and hopefully aids the brain diseases analysis.

# Chapter 6

# Conclusion and Future Work

## 6.1 Conclusion

Over the past decade, researchers have developed various approaches to accurately detect and localize activation of the brain from subtle changes seen in functional MR images. We addressed the challenges of exploring functional brain images for understanding the brain function as well as for the diagnosis and treatment of brain disease [155]. Brain disorders such as: dyslexia, Alzheimer's diseases, Parkinson's diseases, and schizophrenia are common mental diseases disrupting human living. If the researchers can understand the organization and interconnections of these networks, and are in a position to control or manipulate network activities; then researchers are in a position to correct aberrant network activity, and thus treat various mental disorders [156].

A novel method was proposed by using Bayesian networks to learn the structure of effective connectivity among brain regions involved in a functional MR experiment in a data-driven manner. The approach is exploratory in the sense that, unlike previous approaches, no prior model is needed for connectivity analysis. The proposed

approach renders the interactions among brain regions with probability densities, and allows analysis and simulation of dysfunction of neural systems. So far this is the first attempt to analyze the brain effective connectivity in a completely exploratory manner, leading to an attractive new way of interpreting brain function and disease.

The contributions of this report could be summarized in the following four aspects which are mainly involved in brain connectivity analysis with fMRI.

1. The proposed technique is the first method for exploratory analysis of effective connectivity. The conditional probabilities that render the interactions among brain regions in Bayesian networks represent the connectivities in the complete statistical sense. The present method is applicable even when the number of regions involved in the cognitive network is large or unknown. In the functional brain networks, the nodes represent the activated brain regions and a connection between two regions represents an interaction between them. The Maximum A Posteriori (MAP) estimation of the structure of the functional network is derived from fMRI data to maximize the Bayesian Information Criterion (BIC) by using a greedy search algorithm. The present method was tested and compared with other methods using synthetic data, and was demonstrated on fMRI data collected in silent word reading task. In the experiments with synthetic data, the present method was robust to random noise and outperformed SEM in determining the structure. As the number of region increase, the search could fall into a local minima. This problem can be mitigated if a priori knowledge (even if incomplete) of the regions of activation or their connectivity is available. A compromise between confirmatory and exploratory approaches might be more appropriate for brain connectivity analysis. The networks derived from the

present method for silent reading was consistent with the literature, providing a partial validation of the present approach, though a gold standard of the networks of the tasks considered is unavailable.

2. Past studies that have investigated the neuroanatomical representation of English at the levels of orthography, phonology and semantics are reviewed and discussed. However, these neuroimaging studies do not describe how the identified brain regions interact. The present method was applied on fMRI data collected in letter searching task to introduced and discuss a plausible model explaining how these regions are connected at the three subcomponents of language processing. FMRI data was collected in a letter searching task, which was designed to have a cross comparison among reading of high frequency word (HFW), low frequency word (LFW), legal non-word (LNW) and illegal non-word (INW), that will probably reveal the differences between the three components of language processing. The neural systems derived from these tasks by using the present method were compared and validated with the existing literatures to finally conclude a possible general language system. The proposed method of exploring global neural systems from functional imaging data provides an alternate method to study brain function in terms of networks.

3. The method was extended to illustrate its application into brain lesion study through connectivity analysis. The earlier methods of disconnectivity detection are mostly based on comparing the activation patterns. However, there are diseases that are due to disconnections. Author demonstrated how the networks derived from patient and healthy participants using Bayesian network can be used to explore disconnectivities of brain disease. The basic idea is that since Bayesian networks can be used to learn the brain structures from functional data, the networks derived from a

patient and a normal control group doing the same tasks should show the differences of connectivity if any. Hence the functional disorders could be diagnosed. This is a holistic approach to disconnectivity analysis, which benefitted from the characteristic of the complete statistical representation of brain function. The present approach was demonstrated by applying the method on fMRI datasets collected in a counting Stroop task. A brain disorder may due to just a weak connection instead of a complete disconnection. Author extended his method by estimating the parameters of connection for lesion study. The conditional probability that reveals how much a region's activation depends on the others can be computed, and used to represent the strength of an interaction. The approach was demonstrated in a case study of stroke patient.

## 6.2 Future Work

### 6.2.1 Improvements in Bayesian Networks on fMRI

The use of exploratory analysis of brain connectivity gives a new way of studying brain function and disease. However, there are still some improvements could be made to the present method employing Baysian networks.

As a matter of fact, results of the structure searching are not the same as from different searching algorithms. MCMC algorithm was the one that gained the best performance for these experiments. However, after searching, the MCMC algorithm actually returns a sample of structures with scored values. Naturally, the structure with the highest score was chosen. While this is not likely to be the unique choice, sometimes structures with lower scores seem more reasonable with the anatomical

constraints due to the local minima problem as mentioned in the discussion section of Chapter 3. This is a tricky situation as the goal of the method is the exploratory extraction of the structure from data, while results should not go too far from anatomical constraints. This situation suggests that graphical models are more appropriate for lesion study, since whatever structure returned, the destination is there to find the connectivity distinct between patients and healthy controls doing the same tasks and same initial conditions. The structures are supposed to be the same for all healthy controls doing the same task, the difference in structures between patients and normals should show the root of the disorders. For more general structure learning from fMRI data, searching algorithm and score functions should be improved to get more affirmable results.

### 6.2.2 Study Neural Systems

The study of neural network and its interactions is important for relating cognitive theories with brain operations. It is highly unlikely that psychological function is modulated by a single brain construct. On the other hand, it is also unlikely that a single brain region takes part in only one cognitive function. There may not be a single brain area that represents "attention"for instance, but it is more likely that numerous brain areas are involved, whose interactions represent attention operations. The important point is that it may be possible for parts of the same anatomical network to be involved in another function when the interactions change. These notions about brain networks may suggest the reasons why some separate functions cannot occur simultaneously on human beings, since certain brain area is not allowed to effect in two different neural systems at the same time.

Therefore, it is promising if distinct neural systems corresponding to separate brain functions can be detected, which consists of remote or overlapping brain areas. Using the Bayesian networks, the structures can be derived directly from fMRI data scanned from any functional tasks. By studying the structures, a general outline of different neural systems can be learned. So far a possible language network based on the present method has been introduced. However this is far from enough, there are still many other systems to be explored. Even the language network being proposed is ground accepted, it was only representing a tiny aspect of the whole language system. In future research, more fMRI experiments could be designed and performed for more specific cognitive tasks. With the developing of the method, a general frame work of introducing neural system could be built with the assists of other contributions in this area.

### 6.2.3 Simulate Lesion Studies

Functional MRI is increasingly being introduced into clinical practice. Applying the present techniques into lesion studies will lead to a new way of looking into the abnormal brain. Mimicking abnormal neural systems will contribute to investigating the pathological mechanism-underlying mental illness.

The implementation of Bayesian networks in brain disorder analysis has been described in Chapter 5. As the fMRI scan of the patients will normally go with the normal control groups, the idea that the learning results from a patients and a normal control group doing the same tasks should show how the connectivity differs to each other, is thus, feasible. Author will further interpret the abnormal behavior from more particular brain diseases such as dyslexia. As mentioned before, in a structure

each region is responsible for a specific function or behavior. The abnormality at a structure as an event with a probability can be simulated; so that the probabilities of activation of other structures can be predicted. The approach of mimicking lesion study is not completed yet. We will continue this promising research and try to achieve the goal of proposing this method as a general framework to analyze a wide range of brain disorders.

### 6.2.4 Study Hypotheses Made in Brain Disorders

Although many efforts have been done, brain disorders such as dyslexia, Alzheimer's disease, Parkinson's disease and schizophrenia still remain as hypotheses. The research can be carried on to study these hypotheses by implementing graphical models with fMRI data from patients.

Dyslexia is one of the main forms of reading disorders observed when the previously competent reading activity of an adult is disrupted by brain injury or disease. The functional disconnection of the left angular gyrus, which is involved in normal reading task, has been found [102, 157]. Future research will be aimed to test this argument by detecting the neural system involved in reading tasks. From this finding, it may be possible to relate other seemingly disparate syndromes with the reading-related problem, where the same part of brain area belonging to the corresponding neural system is injured or disconnected.

### 6.2.5 Theoretical Work and Validation

Conditional probabilities used in Bayesian networks describe complete behavior of the network, theoretically; covariances describe only second order behavior. Further

investigation of advantages of graphical model implementation will be instructive to the implementation of graphical model with functional imaging.

Independent cliques of probability densities to define independent neural systems will be proposed, by decomposing (factoring) probability densities into marginal densities. When there are many regions showing activities, some of them might be performing the same function as independent neural systems. Doing independent clique analysis will simplify the model and assist independent neural system analysis.

# Bibliography

[1] A.R. McIntosh and F. Gonzalez-Lima. Structural Equation Modeling and Its Application to Network Analysis in Functional Brain Imaging. *Human Brain Mapping*, 2:2–22, 1994.

[2] R. Goebela, A. Roebroecka, D. Kimb, and E. Formisanoa. Investigating directed cortical interactions in time-resolved fmri data using vector autoregressive modeling and granger causality mapping. *Magnetic Resonance Imaging*, 21:1251–1261, 2003.

[3] K. J. Friston. Dynamic causal modelling. *Neuroimage*, 19:1273–1302, 2003.

[4] K. J. Friston, A. P. Holmes, K. J. Worsley, J. B. Poline, C. D. Frith, and R. S. J. Frackowiak. Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping*, 2:189–210, 1995.

[5] K. K. Kwong. Functional magnetic resonance imaging with echo planar imaging. *em Magn. Reson. Q.*, 11:1–20, 1995.

[6] J. Xiong, J.-H. Gao, J. L. Lancaster, and P. T. Fox. Assessment and optimization of functional mri analyses. *Human Brain Mapping*, 4:153–167, 1996.

[7] K. J. Friston, K. J. Worsley, R. S. J. Frackowiak, J. C. Mazziotta, and A. C. Evans. Assessing the significance of focal activations using their spatial extent. *Huamn Brain Mapping*, 1:213–220, 1994.

[8] M. Fitzgerald W. F. Eddy M. A. Mintun D. C. Noll S. D. Forman, J. D. Cohen. Improved assessment of significant activation in functional magnetic resonance imaging (fmri): Use of a cluster-size threshold. *Magn. Reson. Med.*, 33:636–647, 1995.

[9] A. C. Evans K. J. Friston J.-B. Poline, K. J. Worsley. Combining spatial extent and peak intensity to test for activations in functional imaging. *Neuroimage*, 5:83–96, 1997.

[10] K. J. Worsley and K. J. Friston. Analysis of fMRI time-series revisited - again. *Neuroimage*, 2:173–181, 1995.

[11] J. C. Rajapakse and J. Piyaratna. Probabilistic framework to segmentation of statistical parametric maps. *IEEE Trans. Biomedical Engineering*, 48:1186–1194, 2001.

[12] Y. Wang and J. C. Rajapakse. Contextual modeling of functional mr images with conditional random fields. *IEEE Transactions on Medical Imaging*, 25:no. 6, 2006.

[13] A. H. Andersen, D. M. Gash, and M. J. Avison. Principal component analysis of the dynamic response measured by fmri: A generalized linear system framework. *Magn. Reson. Imag.*, 17(6):795–815, 1999.

[14] M. J. McKeown, S. Makeig, G. G. Brown, T. P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski. Analysis of fmri data by blind separation into independent spatial components. *Human Brain Mapping*, 6:160–188, 1998.

[15] M. J. Fadili, S. Ruan, D. Bloyet, and B. Mazoyer. Unsupervised fuzzy clustering analysis of fmri series. *proceedings of the 20th Annual International Conf. of IEEE Eng. in Med. Biol.*, 20(2):696–699, 1998.

[16] Hamid Soltanian-Zadeh, Gholam-Ali Hossein-Zadeh, and Babak A. Ardekani. fmri activation detection in wavelet signal subspace. *MI 48202, USA*.

[17] W. D. Penny, K.E. Stephan, A. Mechelli, and K.J. Friston. Modelling functional integration: A comparison of structural equation and dynamic causal models. *NeuroImage*, 23:264–274, 2004a.

[18] Michael I. Jordan. *Learning in Graphical Models*. The MIT Press, 1999.

[19] D. Heckerman. A tutorial on learning with bayesian networks. *Technical Report MSR-TR-95-06, Microsoft Research, Advanced Technology Division*, March 1995.

[20] J. Wesley Barnes. *Statistical Analysis for Engineers and Scientists*. The McGraw-Hill, Inc., 1994.

[21] E. Paulesu, U. Frith, M. Snowling, A. Gallagher, J. Morton, R.S.J. Frackowiak, and C.D. Frith. Is developmental dyslexia a disconnection syndrome? evidence from pet scanning. *Brain*, 119:143–157, 1996.

[22] K.R. Pugh, W.E. Mencl, B.A. Shaywitz, S.E. Shaywitz, R.K. Fulbright, R.T. Constable, P. Skudlarski, K.E Marchione, A.R. Jenner, J.M. Flectcher, A.M. Liberman, D.P. Shankweiler, L. Katz, C. Lacadie, and J.C. Gore. The angular gyrus in developmental dyslexia: Task-specific differences in functional connectivity within posterior cortex. *American Psychological Society*, 11(1), 2000.

[23] D.E. Welchew, C. Ashwin, K. Berkouk, R. Salvador, J. Suckling, S. Baron-cohen, and E. Bullmore. Functional disconnectivity of the medial temporal lobe in asperger's syndrome. *Biological Psychiatry*, 57:991–998, 2005.

[24] H.C. Whalley, E. Simonotto, I. Marshall, D.G.C. Owens, N.H. Goddard, E.C. Johnstone, and S.M. Lawrie. Functional disconnectivity in subjects at high genetic risk of schizophrenia. *Brain*, 10, 2005.

[25] X. Zheng and J. C. Rajapakse. Learning functional strucutre from fmri images. *Neuroimage*, in press, 2006.

[26] L. Tamm, V. Menon, C.K. Johnston, D.R. Hessl, and A.L. Reiss. fmri study of cognitive interference processing in females with fragile x syndrome. *Journal of Cognitive Neuroscience*, 14(2):160–171, 2002.

[27] N. C. Andreasen. *Brave New Brain, Conquering Mental Illness in the Era of the Genome.* Oxford University Press, 2001.

[28] Department of Kinesiology and Health Education University of Texas at Austin. Brain. *http://www.edb.utexas.edu/syllabus/farrar/lectures/brain.html*, 1999.

[29] R.S.J. Frackowiak, K.J. Friston, C.D. Frith, R.J. Dolan, and Mazziotta J.C. *Human Brain Function.* Academic Press, San Diego, 1997.

[30] A. Parry and P.M. Matthews. Functional magnetic resonance imaging (fmri): A window into the brain. *University of Oxford, The John Radcliffe Hospital*, 2002.

[31] G. Muller. Introduction to medical imaging case study. *Embedded Systems Institute - The Netherlands.*

[32] P. A. Bandettini, A. Jesmanowicz, E.C. Wong, and J. S. Hyde. Processing strategies for time-course data sets in functional MRI of the human brain. *Magn Reson Med*, 30:161–173, 1993.

[33] J. C. Rajapakse, F. Kruggel, J. M. Maisog, and D. Y. Cramon. Modeling hemodynamic response for analysis of functional MRI time-series. *Human Brain Mapping*, 6:283–300, 1998.

[34] UK Wellcome Department of Cognitive Neurology, London. Statistical Parametric Mapping. *http://www.fil.ion.ucl.ac.uk/spm/*, 2004.

[35] K. J. Friston, J. Ashburner, C.D. Frith, J-B Poline, J.D. Heather, and R.S.J. Frackowiak. Spatial registration and normalization of images. *Huamn Brain Mapping*, 2:165–189, 1995.

[36] J. Ashburner and K. Friston. Voxel-based morphometry - the methods. *NeuroImage*, 11:805–821, 2000.

[37] K. J. Friston, S. Williams, R. Howard, R. S. Frackowiak, and R. Turner. Movement-related effect in fMRI time-series. *Magn Reson Med*, 35:346–355, 1996.

[38] J.L.R. Andersson, C. Hutton, J. Ashburner, R. Turner, and K. Friston. Modelling geometric deformations in epi time series. *NeuroImage*, 13:903–919, 2001.

[39] J. Ashburner and K. Friston. Multimodal image coregistration and partitioning - a unified framework. *NeuroImage*, 6(3):209–217, 1997.

[40] P.T. Fox. Spatial normalization origins: Objectives, applications, and alternatives. *Human Brain Mapping*, 3:161–164, 1995.

[41] K.J. Friston. Functional connectivity: the principal-component analysis of large (pet) data sets. *Journal of Cerebral Blood Flow and Metabolism*, 13:5–14, 1993.

[42] K.J. Friston. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2:56–78, 1994.

[43] Yang Kanyan. Higher-order brain connectivity. *First Year Report of Ph.D, Nanyang Technological University*, 2003.

[44] S.L. Bressler and J.A.S. Kelso. Cortical coordination dynamics and cognition. *Trends in Cognitive Sciences*, 5:26–36, 2001.

[45] Olaf Sporns. Network Analysis, Complexity, and Brain Function. *Complexity.*

[46] Christian Buchel and K. J. Friston. Effective connectivity and neuroimaging. *SPM course notes*, 1997.

[47] Ed Bullmore, Barry Horwitz, Garry Honey, Mick Brammer, Steve Williams, and Tonmoy Sharma. How Good is Good Enough in Path Analysis of fMRI Data? *NeuroImage*, 11:289–301, 2000.

[48] G. D. Honey, C. H. Y. Fu, Mick Brammer, T. J. Croudace, J. Suckling, E. M. Pich, Steve Williams, and Ed Bullmore. Effects of Verbal Working Memory Load on Corticocortial Connectivity Modeled by Path Analysis of Functional Magnetic Resonace Imaging Data. *NeuroImage*, 17:573–582, 2002.

[49] O. Sporns and G. Tononi. Classes of network connectivity and dynamics. *Complexity*, 7:28–38, 2002.

[50] Christian Buchel and K. J. Friston. Dynamic Changes in Effective Connectivity Characterized by Variable Parameter Regression and Kalman Filtering. *Human Brain Mapping*, 6:403–408, 1998.

[51] A.R. McIntosh, C. L. Grady, L. G. Ungerieider, J. V. Haxby, S. I. Rapoport, and B. Horwitz. Network Analysis of Cortial Visual Pathways Mapped with PET. *The Journal of Neuroscience*, 14(2):655–666, 1994.

[52] K. J. Worsley, J. Chen, J. Lerch, and A.C. Evans. Comparing functional connectivity via thresholding correlations and singular value decomposition. *Phil. Trans. R. Soc. B.*, 360:913–920, 2005.

[53] J. Cao and K.J. Worsley. The geometry of correlation fields, with an application to functional connectivity of the brain. *Annals of Applied Probability*, 9:1021–1057, 1999.

[54] R. Baumgartner, L. Ryner, W. Richter, R. Summers, M. Jarmasz, and R. Somorjai. Comparison of two exploratory data analysis methods for fmri: fuzzy

clustering vs. principal component analysis. *Magnetic Resonance Imaging*, 18:89–94, 2000.

[55] V. G. van de Ven, E. Formisano, D. Prvulovic, C. H. Roeder, and D. E. J. Linden. Functional connectivity as revealed by spatial independent component of fmri during rest. *Human Brain Mapping*, 22:165–178, 2004.

[56] D. Cordes, V. Haughton, J. D. Carew, K. Arfanakis, J. D. Crew, and K Maravilla. Hierarchical clustering to measure connectivity in fmri resting state data. *Maganetic Resonance Imaging*, 20:305–317, 2002.

[57] M.S. Goncalves and D.A. Hall. Connectivity analysis with structural equation modelling: an example of the effects of voxel selection. *Neuroimage*, 20:1455–1467, 2003.

[58] S. Dodel, N. Golestani, C. Pallier, V. ElKouby, D. L. Bihan, and J. Poline. Condition-dependent functional connectivity: syntax networks in bilinguals. *Phil. Trans. R. Soc. B.*, 360:921–935, 2005.

[59] R. Salvador, J. Suckling, C. Schwarzbauer, and E. Bullmore. Undirected graphs of frequency-dependent functional connectivity in whole brain networks. *Phil. Trans. R. Soc. B.*, 360:937–946, 2005.

[60] Kenneth A. Bollen. *Structural Equations with Latent Variables*. John Wiley and Sons, Inc., 1989.

[61] N.K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann. Neurophysiological investigation of the basis of the fmri signal. *Nature*, 412:150–157, 2001.

[62] J. Geweke. Measurement of linear dependence and feedback between multiple time series. *J Am Stat Assoc*, 77:304–313, 1982.

[63] K.J. Friston. Bayesian estimation of dynamical systems: an application to fmri. *Neuroimage*, 16:513–530, 2002.

[64] B. J. Krause, B. Horwitz, J. G. Taylor, D. Schmidt, F. M. Mottaghy, H. Herzog, U. Halsband, and H. W. Muller-Gartner. Network Analysis in Episodia Encoding and Retrieval of Word-pair Associates: A PET Study. *The Journal of Neuroscience*, 11:3293–3301, 1999.

[65] Lars Nyberg, A. R. McIntosh, Roberto Cabeza, Lars-Goran Nilsson, Sylvain Houle, Reza Habib, and Endel Tulving. Network Analysis of Positron Emission Tomography Regional Cerebral Blood Flow Data: Ensemble Inhibition during Episodic Memory Retrieval. *The Journal of Neuroscience*, 16(11):3753–3759, 1996.

[66] D. Bavelier, A. Tomann, C. Hutton, T. Mitchell, D. Corina, G. Liu, and H. Neville. Visual Attention to the Periphery is Enhanced in Congenitally Deaf Individuals. *The Journal of Neuroscience*, 20:RC93 1–6, 2000.

[67] G.D. Honey, C.H.Y. Fu, J. Kim, M.J. Brammer, T.J. Groudace, J. Suckling, E.M. Pich, S.C.R. William, and E.T. Bullmore. Effects of Verbal Working Memory Load on Corticocortial Connectivity Modeled by Path Analysis of Functional Magnetic Resonace Imaging Data. *NeuroImage*, 17:573–582, 2002.

[68] Reza Nezafat, Reza Shadmehr, and Henry H. Holcomb. Long-term Adaptation to Dynamics of Reading Movement: A PET Study. *Exp Brain Res*, 140:66–76, 2001.

[69] K.A. McKiernan, L.L. Conant, A. Chen, and J.R. Binder. Development and cross-validation of a model of linguistic processing using neural network and path analyses with fmri data. *NeuroImage*, 13(6):2–2, 2001.

[70] K.L. Petersson, A. Reis, S. Askelof, A. Castro-Caldas, and M. Ingvar. Language processing modulated by literacy: A network analysis of verbal repetition in literate and illiterate subjects. *Congnitive Neuroscience*, pages 364–382, 2000.

[71] A. Mechelli, W.D. Penny, C.J. Price, D.R. Gitelman, and K.J. Friston. Effective connectivity and intersubject variability: Using a multisubject network to test differences and commonalities. *NeuroImage*, 17:1459–1469, 2002.

[72] M. Eichler. A graphical approach for evaluating effective connectivity in neural systems. *Phil. Trans. R. Soc. B.*, 360:953–967, 2005.

[73] P. A. Valdes-Sosa, J. M. Sanchez-Bornot, A. Lage-Castellanos, M. Vega-Hernandez, J. Bosch-Bayard, L. Melie-Garcia, and E. Canales-Rodriguez. Estimating brain functional connectivity with sparse multivariate autoregression. *Phil. Trans. R. Soc. B.*, 360:969–981, 2005.

[74] W. D. Penny, K.E. Stephan, A. Mechelli, and K.J. Friston. Comparing dynamic causal models. *NeuroImage*, 22:1157–1172, 2004b.

[75] W.D. Penny, Z. Ghahramani, and K.J. Friston. Bilinear dynamical systems. *Phil. Trans. R. Soc. B*, 360:983–993, 2005.

[76] J. C. Rajapakse, C. L. Tan, X. Zheng, S. Mukhopadhyay, and Kanyan Yang. Exploratory analysis of brain connectivity with ica. *IEEE Engineering in Medicine and Biology Magazine*, 25(2):102–111, 2006.

[77] Carvin R. Maurer and J. Michael Fitzpatrik. A review of medical image registration. *American Association of Neurological Surgeon*, pages 17–44, 1993.

[78] J. B. Antoine Maintz and Max A. Viergever. An overview of medical image registration methods. *Imaging Science Department, Imaging Center Utrecht.*

[79] F. V. Jensen. *Bayesian Networks and Decision Diagrams.* Springer, 2001.

[80] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. *Probabilistic Networks and Expert Systems.* Springer, 1999.

[81] K. P. Murphy. Dynamic bayesian networks: Representation, inference and learning. *PhD thesis, UC Berkeley, Computer Science Division, http://www.ai.mit.edu/ murphyk/Thesis/thesis.html*, 2002.

[82] D. Heckerman and D. Geiger. Likelihoods and parameter priors for bayesian networks. *MicroSoft Research, Redmond, WA.*, Technical Report:MSR–TR–95–54, 1995.

[83] R. Kass and A. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90:773–795, 1995.

[84] D.Margaritis. Learning bayesian network model structure from data. *Thesis, School of Computer Science, Carnegie Mellon University, Pittsburgh.*

[85] Dartmouth College. The fMRI Data Center. fmridc. *http://www.fmridc.org.*

[86] K. P. Murphy. The Bayes Net Toolbox for Matlab. *http://www.ai.mit.edu/ murphyk/Software/BNT/usage.html.*

[87] Andrea Mechelli, Karl J. Friston, and Cathy J. Price. The Effects of Presentation Rate During Word and Pseudoword Read: A Comparison of PET and fMRI. *Cognitive Neuroscience*, 12:145 – 156, 2000.

[88] K. Friston, A.P. Holmes, and K.J. Worsley. How many subjects constitute a study? *Neuroimage*, 10:1–5, 1999.

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

[90] K. J. Worsley, J.E. Taylor, F. Tomaiuolo, and J. Lerch. Unified univariate and multivariate random field theory. *Neuroimage*, 23:189–195, 2004.

[91] Jack L.Lancaster, Peter T. Fox, Shawn Mikiten, and Lacy Rainey. Talairach Deamon Database.

[92] Jean Talairach and Pierre Tournoux. *Co-planar Stereotaxic Atlas of the Human Brain.* Thieme Medical Publishers, Inc., 1988.

[93] M. Brett. The mni brain and the talairach atlas. *http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml*, 2002.

[94] K.H.S. Kim, N.R. Relkin, K. Lee, and J. Hirsch. Distinct cortical areas associated with native andsecond languages. *Nature*, 388:171–174, 1997.

[95] L.H. Tan, J.A. Spinks, J. Gao, H. Liu, C.A. Perfetti, J.H. Xiong, K.A. Stofer, Y. Pu, Y. Liu, and P.T. Fox. Brain activation in the processing of chinese characters and words: A functional mri study. *Human Brain Mapping*, 10:16–27, 2000.

[96] B. Kolb and I.Q. Whishaw. *Fundamantal of Human Neuropsychology.* W. H. Freman and Company, 1996.

[97] BrainPlace.com. Brain spect information and resources. brain systems, functions and problems. *http://www.brainplace.com/bp/brainsystem/*, 2005.

[98] C.J. Fiebach, A.D. Friederici, K. Muller, and D.Y.V. Cramon. fmri evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience*, 14(1):11–23, 2002.

[99] C.J. Price. The anatomy of language: contributions from functional neuroimaging. *J. Anat.*, 179:335–359, 2000.

[100] M. Hampson, B. S. Peterson, P. Skudlarski, J. C. Gatenby, and J. C. Gore. Detection of functional connectivity using temporal correlations in MR images. *Human Brain Mapping*, 15:247–262, 2002.

[101] B. Horwitz and A.R. Braun. Brain network interactions in auditory, visual and linguistic processing. *Brain and Language*, 89(2):377–384, 2004.

[102] B. Howitz, J. M. Rumsey, and B. C. Donohue. Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proc. Natl. Acad. Sci. USA*, 95:8939–8944, 1998.

[103] R. Matsumoto, D.R. Nair, E. LaPresto, I. Najm, W. Bingaman, H. Shibasaki, and H.O. Luders. Functional connectivity in the human language system: a cortico-cortical evoked potential study. *Brain*, pages 2316–2330, 2004.

[104] Geoffrey M. Maruyama. *Basics of Structural Equations Modeling.* John Wiley and Sons, Inc., 1989.

[105] J.R. Binder. Functional magnetic resonance imaging of language cortex. *International Journal of Imaging Systems and Technology*, 6:280–288, 1995.

[106] E. R. Kandel. Essentials of neural sciences and behaviour. *Norwalk, Appleton, and Lange*, pages 633–649, 1995.

[107] C. Buchel, C. Frith, and K. Friston. *Functional integration: methods for assessing interactions among neuronal systems using brain imaging.* 1999.

[108] M. Brown and P. Hagoort. *The cognitive neuroscience of language challenges and future directions.* 1999.

[109] Bryan Kolb and Ian Q. Whishaw. *Fundamantal of Human Neuropsychology.* W. H. Freman and Company, 1996.

[110] Jeffrey R. Binder, Julie A. Frost, Thomas A. Hammeke, Robert W. Cox, Stephen M. Rao, and Thomas Prieto. Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience*, 17(1):353–362, 1997.

[111] David Caplan. Language and the brain. *The Harvard Mahoney Neuroscience Institute Letter*, 4:4, 1995.

[112] ICBM. International consortium for brain mapping. *http://www.loni.ucla.edu/ICBM*.

[113] Matthew Brett. mni2tal. *http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml*.

[114] B. Faw. Pre-frontal executive committee for perception, working memory, attention, long-term memory, motor control, and thinking: A tutorial review. *Consciousness and Cognition*, 12):83–139, 2002.

[115] T. Wu, K. Kansaku, and M. Hallett. How self-initiated memorized movements become automatic: A functional mri study. *Neurophysiol*, 91:1690–1698, 2004.

[116] A.G. He, L.H. Tan, Y. Tang, G.A. James, P. Wright, M.A. Eckert, P.T. Fox, and Y. Liu. Modulation of neural connectivity during tongue movement and reading. *Human Brain Mapping*, 18:222–232, 2003.

[117] J. Joseph, K. Noble, and G. Eden. The neurobiological basis of reading. *Journal of Learning Disabilities*, 34:566–579, 2001.

[118] D. Howard, K. Patterson, R. Wise, W. D. Brown, K. Friston, C. Weiller, and R. Frackowiak. The cortical localization of the lexicons. *Brain*, 115:1769–1782, 1992.

[119] S. E. Petersen, P.T. Fox, M.I. Posner, M. Mintun, and M.E. Raichle. Positron emission tomographic studies of cortical anatomy of single-word processing. *Nature*, pages 585–589, 1988.

[120] S. E. Petersen, P.T. Fox, M.I. Posner, M. Mintun, and M.E. Raichle. Positron emission tomographic studies of the processing of single words. *Journal of Cognitive Neuroscience*, pages 153–170, 1989.

[121] S. E. Petersen, P.T. Fox, A.Z. Snyder, and M.E. Raichle. Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. *Science*, pages 1041–11044, 1990.

[122] S.L. Small, D.C. Noll, C.A. Perfetti, P. Hlustik, R. Wellington, and W. Schneider. Localizing the lexicon for reading aloud: replication of a pet study using fmri. *NeuroReport*, 7:961–965, 1996.

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

[124] L. Friedman, J.T. Kenny, A.L. Wise, D. Wu, T.A. Stuve, D.A. Miller, J.A. Jesberger, and J.S. Lewin. Brain activation during silent word generation evaluated with functional mri. *Brain and Language*, 64:231–256, 1998.

[125] E. Paulesu, B. Goldacre, P. Scifo, S.F. Cappa, M.C. Gilardi, D.P. Castiglioni, and F. Fazio. Functional heterogeneity of left inferior frontal cortex as revealed by fmri. *NeuroReport*, 8:2011–2016, 1997.

[126] L. Rueckert, I. Appollonio, J. Grafman, P. Jezzard, R.J. Johnson, D. Le Bihan, and R. Turner. Magnetic resonance imaging functional activation of left frontal cortex during covert word production. *Journal of Neuroimaging*, 4:67–70, 1994.

[127] K.R. Pugh, B.A. Shaywitz, S.E. Shaywitz, R.T. Constable, P. Skudlarski, R.K. Fulbright, R.A. Bronen, D.P. Shankweiler, L. Katz, J.M. Flectcher, and J.C. Gore. Cerebral organization of component processes in reading. *Brain*, 119:1221–1238, 1996.

[128] B.A. Shaywitz, K.R. Pugh, R.T. Constable, S.E. Shaywitz, R.A. Bronen, P.K. Fulbright, D.P. Shankweiler, L. Katz, J.M. Fletcher, P. Skudlarski, and J.C. Gore. Localization of semantic processing using functional magnetic resonance imaging. *Human Brain Mapping*, 2:149–158, 1995.

[129] R.L. Buckner and S.E. Petersen. What has neuroimaging told us about prefrontal cortex involvement in long-term memory retrieval? *Seminars in the Neurosciences*, 8:47–55, 1996.

[130] R.L. Buckner, S.E. Petersen, and M.E. Raichle. Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. *Journal of Neurophysiology*, 74:2163–2173, 1995.

[131] J.B. Demb, J.E. Desmond, A.D. Wagner, C.J. Vaidya, G.H. Glover, and J.D.E. Gabrieli. Semantic encoding and retrieval in the left inferior prefrontal cortex: a functional mri study of task difficulty and process specificity. *The Journal of Neuroscience*, 15(9):5870–5878, 1995.

[132] J.D.E. Gabrieli, J.E. Desmond, J.B. Demb, A.D. Wagner, M.V. Stone, C.J. Vaidya, and G.H. Glover. Functional magnetic resonance imaging of semantic memory processes in the frontal lobes. *Psychological Science*, 7(5):278–283, 1996.

[133] J. Sergent, E. Zuck, M. Levesque, and B. MacDonald. Positron emission tomography study of letter and object processing: empirical findings and methodological considerations. *Cerebral Cortex*, 2:68–80, 1992.

[134] J.R. Binder, S.M. Rao, T.A. Hammeke, F.Z. Yetkin, A. Jesmanowicz, P.A. Bandettini, E.C. Wong, L.D. Estkowski, M.D. Goldstein, V.M. Haughton, and J.S. Hyde. Functional magnetic resonance imaging of human auditory cortex. *Annals of Neurology*, 35(6):662–672, 1994.

[135] C.J. Price, R.J.S. Wise, E.A. Warburton, C.J. Moore, D. Howard, K. Patterson, R.S.J. Frackowiak, and K.J. Friston. Hearing and saying: The functional neuro-anatomy of auditory word processing. *Brain*, 119:919–931, 1996.

[136] E. Warburton, R.J. Wise, C.J. Price, C. Weiller, U. Hadar S. Ramsay, and R.S. Frackowiak. Noun and verb retrieval by normal subjects. studies with pet. *Brain*, 119:159–179, 1996.

[137] J.F. Demonet, F. Chollet, S. Ramsay, D. Cardebat, J.L. Nespoulous, R. Wise, A. Rascol, and R. Frackowiak. The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115:1753–1768, 1992.

[138] J.F. Demonet, C. Price, R. Wise, and R.S.J. Frackowiak. Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: A positron-emission tomography study in normal human subjects. *Neuroscience Letters*, 182:25–28, 1994.

[139] A.N. Herbster, M.A. Mintun, R.D. Nebes, and J.T. Becker. Regional cerebral flow during word and nonword reading. *Human Brain Mapping*, 5:84–92, 1997.

[140] J.M. Rumsey, B. Horwitz, B.C. Donohue, K. Nace, and J.M. Maisog P. Andreason. Phonologic and orthographic components of word recognition: A pet-rcbf study. *Brain*, 120:739–759, 1997.

[141] J.A. Fiez and S.E. Petersen. Pet as part of an interdisciplinary approach to understanding processes involved in reading. *Psychological Science*, 4:287–292, 1993.

[142] G. McCarthy, A.M. Blamire, D.L. Rothman, R. Gruetter, and R.G. Shulman. Echo-planar magnetic resonance imaging studies of frontal cortex activation during word generation in humans. *Proceedings of the National Academy of Sciences USA*, 90:4952–4956, 1993.

[143] C.J. Price, C.J. Moore, G.W. Humphreys, and R.J.S. Wise. Segregating semantic from phonological processes during reading. *Journal of Cognitive Neuroscience*, 9:727–733, 1997.

[144] R. Wise, F. Choller, U. Hadar, K. Friston, E. Hoffner, and R.S.J. Frackowiak. Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain*, 114:1803–1807, 1991.

[145] J.A. Fiez and S.E. Petersen. Neuroimaging studies of word reading. *Proceedings of the National Academy of Science of the United States of America*, 95:914–921, 1998.

[146] G. Bush, P.J. Whalen, B.R. Rosen, M.A. Jenike, S.C. McInerney, and S.L. Rauch. The counting stroop: An interference task specialized for functional neuroimagingvalidation study with functional mri. *Human Brain Mapping*, 6:270–282, 1998.

[147] A.S. Field, Y.F. Yen, J.H. Burdette, and A.D. Elster. False cerebral activation on bold functional mr images: study of low-amplitude motion weakly correlated to stimulus. *AJNR Am J Neuroradiol*, 21(8):1388–1396, 2000.

[148] M. Gavrilescu, G.W. Stuart, A. Waites, G. Jackson, I.D. Svalbe, and G.F. Egan. Changes in effective connectivity models in the presence of task-correlated motion: an fmri study. *Human Brain Mapping*, 21(2):49–63, 2004.

[149] G. Hayward, G.M. Goodwin, C.J. Harmer, and R.W. McCharley. The role of the anterior cingulate cortex in the counting stroop task. *Exp Brain Res*, 154:355–358, 2004.

[150] L.M. Shin, P.J. Whalen, P.K. Pitman, G. Bush, M.L. Macklin, N.B. Lasko, S.P. Orr, S.C. McInerney, and S.L. Rauch. An fmri study of anterior cingulate function in posttraumatic stress disorder. *Society of Biological Psychaitry*, 0006 - 3223(01):012–015, 2001.

[151] D. Vanlancker-Sidtis, A.R. McIntosh, and S. Grafton. Pet activation studies comparing two speech tasks widely used in surgical mapping. *Brain and Language*, 85:245–261, 2003.

[152] D. Heckerman and D. Geiger. Learning gaussian networks. *Tenth Conference on Uncertainty in Artificial Intelligence, Seattle, WA*, July 1994.

[153] Dan Geiger. Learning gaussian networks. *Technical Report, Microsoft Research, Advanced Technology Division, Micorsoft Corporation*, pages MSR–TR–94–10, March 1994.

[154] R. Shachter. Bayesian analysis in expert systems. *Statistical Science*, 8:219–282, 1993.

[155] Jagath C. Rajapakse. Neural Systems Modelling with functional Magnetic Resonance imaging (fmri). *http://www.ntu.edu.sg/home/asjagath/neural-systems-modeling.htm*.

[156] Shawn Mikula. Research on determining the large-scale functional and anatomical organization in the brains of humans and non-human primates. *http://mind-brain.com/researchinterests.html*.

[157] Shally E. Shaywitz and et al. Functional disruption in the organization of the brain for reading in dyslexia. *Proc. Natl. Acad. Sci. USA*, 95:2636–2641, 1998.