Detecting Language Activations with Functional Magnetic Resonance Imaging

Andrea Mechelli Ph. D. Thesis

Wellcome Department of Imaging Neuroscience, Institute of Neurology

University College London

ProQuest Number: U643619

All rights reserved

INFORMATION TO ALL USERS The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest U643619

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code. Microform Edition © ProQuest LLC.

> ProQuest LLC 789 East Eisenhower Parkway P.O. Box 1346 Ann Arbor, MI 48106-1346

Acknowledgements

My sincerest thanks to Cathy Price and Karl Friston. I have benefited enormously, both personally and intellectually, from Cathy's formidable enthusiasm, patience and generous support. I am also much indebted to Karl for his encouragement and enthusiastic supervision. I wish to thank my colleagues at the Functional Imaging Laboratory, in particular Chloe Hutton, Rik Henson, Will Penny, Uta Noppeney and Marilu Gorno Tempini for their advice and friendship.

I am also grateful to my parents and all my family and friends for their support and encouragement. Above all, I wish to thank James.

Andrea Mechelli 4th July 2002

Abstract

This thesis investigates a number of factors that affect sensitivity to language activations in functional Magnetic Resonance Imaging (fMRI). In the first part, I investigate the impact of experimental design parameters on the ability to detect language activations. These parameters include stimulus rate, stimulus duration, stimulus amplitude, epoch length and stimulus ordering. Crucially, they may affect sensitivity in multiple ways that include neurophysiological, efficiency-mediated and BOLD saturation effects. I illustrate and discuss these effects by presenting biophysical simulations and fMRI studies of single word and pseudoword reading. In addition, I focus on the differential effects of the above parameters in Positron Emission Tomography and fMRI studies.

In the second part, I investigate the impact of the analysis used to estimate effects of interest from the data. I compare event-related and epoch analyses and show that, even in the context of blocked design fMRI, an event-related model may provide greater sensitivity than an epoch model. I then address the notion that experimentally-induced effects may be detected not only as task-dependent changes in regional responses but also as changes in connectivity amongst functionally connected regions. These two complementary approaches are motivated by two fundamental principles of brain organisation: functional specialisation and functional integration. I present two fMRI studies investigating the neural correlates of reading words and pseudowords in terms of functional specialisation and functional integration. Furthermore, in both studies I address the issue of inter-subject variability, which may be a critical determinant of sensitivity.

3

Table of Contents

Chapter 1 General Introduction	
1.1 Introduction	13
1.2 Response Sampling	15
1.3 The Effects of Experimental Design in fMRI	18
1.3.1 BOLD Saturation effects	18
1.3.2 Neurophysiological effects	19
1.3.3 Efficiency-mediated effects	20
1.4 The Effects of Experimental Design in PET and fMRI	22
1.4.1 Physiological Parameters measured in PET and fMRI	23
1.4.2 Data Acquisition in PET and fMRI	25
1.4.3 The Nature of the Baseline in PET and fMRI	26
1.5. Epoch vs. Event-Related Analysis	29
1.6 Characterizing Experimentally-Induced Effects as Neuronal Interactions	32
1.6.1 Functional Specialisation and Functional Integration	33
1.6.2 Functional and Effective Connectivity	34
1.6.3 Inferring Effective Connectivity	35
1.7 Summary	37

Chapter 2 Methods: Functional Magnetic Resonance Imaging

2.1 Introduction	
2.2 The physics of NMR	
2.2.1 Spin	40
2.2.2 Net magnetisation	
2.2.3 Radiofrequency magnetic fields	
2.2.4 Relaxation	45
2.2.5 Image formation: frequency and phase encoding	
2.2.6 Voxels	
2.2.7 Image contrast	
2.2.8 Ultrafast MRI sequences: Echo-Planar Imaging	51
2.3 fMRI and the magnetic properties of blood	
2.3.1 BOLD contrast in fMRI	
2.3.2 Neurophysiology and BOLD	53

Chapter 3 Methods: Statistical Analysis of fMRI Time Series

3.1 Introduction	56
3.2 Pre-processing	57
3.1 Realignment	58
3.1a Spatial	. 58
3.1b Temporal	. 58
3.2 Coregistration	. 59
3.3 Spatial normalisation	. 59
3.4 Spatial smoothing	.60
3.3 Characterising haemodynamic responses using the General Linear Model	.61
3.3.1 Parameter estimation using the General Linear Model	.63
3.3.2 Residual (error) variance	.64
3.3.3 T statistics	65
3.3.4 Global Scaling	.66
3.3.5 High Pass Filter	.67
3.3.6 Statistical inference and the Theory of Gaussian Fields	.67
3.3.7 Inferences about subjects and populations: Random vs. Fixed effect analyses	.70
3.4 Characterising Effective Connectivity using Structural Equation Modelling	.71
3.4.1 Functional and Effective Connectivity	.71
3.4.2 Structural Equation Modelling	.72
3.4.3 Model estimation	.72
3.4.4 Moderator variables	.74
3.4.5 Interpretation of path coefficients	.75
3.4.6 Implementation of SEM	.75

Chapter 4 Nonlinear Coupling Between Evoked rCBF and BOLD Signals: A Simulation Study of Haemodynamic Responses

4.1 Introduction	78
4.2 The Haemodynamic Model	78
4.2.1 The balloon component4.2.2 The rCBF component	78 81
4.3 The Biophysical Simulations	82
4.3.1 (i) Haemodynamic Refractoriness	83 84
4.3.3 (iii) Stimulus Presentation Rate and SOA	
4.3.5 (v) The form of the evoked neural response	

4.4 Methods	
4.4.1 Simulating rCBF and BOLD responses	
4.4.2 Estimating average rCBF and BOLD responses	89
4.5 Results	91
4.5.1 (i) Haemodynamic Refractoriness	91
4.5.2 (ii) Epoch Length	93
4.5.3 (iii) Stimulus Presentation Rate and SOA	94
4.5.4 (iv) Stimulus amplitude	95
4.5.5 (v) The form of the evoked neural response	98
4.6 Discussion	100
4.6.1 (i) Haemodynamic Refractoriness	
4.6.2 (ii) Epoch Length	
4.6.3 (iii) Stimulus Presentation Rate and SOA	101
4.6.4 (iv) Stimulus amplitude	
4.6.5 (v) The form of the evoked neural response	
4.7 General Discussion	

Chapter 5 Estimating Efficiency a priori: a Comparison of Blocked and Randomised Designs

5.1 Introduction	107
5.2 Methods	110
5.2.1 Subjects	110
5.2.2 Design	110
5.2.3 Stimuli	112
5.2.4 Scanning Technique	112
5.2.5 Data Analysis	113
5.3 Results	115
4.3.1 Combined Analysis	
4.3.2 Independent Analyses	117
5.4 Discussion	119
5.4.1 Differential Efficiency for Blocked and Randomised Sequences	119
5.4.2 Differential Error Variance for Blocked and Randomised Sequences	120
5.5 Conclusion	121

Chapter 6 The effects of presentation rate during word and pseudoword reading: a comparison of PET and fMRI

6.1 Introduction	
6.2 Methods	124
6.2.1 Subjects	124
6.2.2 Design and Task	124
6.2.3 Stimuli	
6.2.4 fMRI Scanning Technique	
6.2.5 Data Analysis	
6.3 Results	128
6.3.1 Effects of Reading relative to Fixation	
6.3.2 The Effects of Rate Common to Words and Pseudowords	
6.3.3 Interactions between Stimulus Rate and Word Type	
6.3.4 Summary of Results	
6.4 Discussion	134
6.4.1 Positive Linear Effect of Rate	
6.4.2 Negative Linear effects of rate	136
6.5 Conclusion	139

Chapter 7 Comparing event-related and epoch analysis in blocked design fMRI

7.1 Introduction	140
7.2 Methods	142
7.2.1 Analysis 1: Event-Related and Epoch Regressors	143
7.2.2 Analysis 2: Event-Related, Epoch and First-stimulus Regressors	143
7.2.3 Analysis 3: Event-Related, Epoch and Last-stimulus Regressors	144
7.3 Results	144
7.3.1 Analysis 1	144
7.3.2 Analysis 2 & 3	146
7.4 Discussion	147
7.5 Conclusion	148

Chapter 8 Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies and limitations

8.1 Introduction	149
8.1.1 Dual Route and Connectionist Models of Reading8.1.2 Functional Imaging Studies of Word and Pseudoword Reading	149 151
8.2 Methods	154
 8.2.1 Subjects	154 154 155 155 155
8.3 Results	157
8.3.1 Common Effects for Word and Pseudoword Reading8.3.2 Word type Effects8.3.3 Duration Effects	157 158 160
8.4 Discussion	162
8.4.1 Main findings8.4.2 Implications for cognitive models of reading	162 164
8.5 Conclusions	167

Chapter 9 Effective Connectivity and Inter-Subject Variability: using a Multi-Subject Network to test Differences and Commonalities

9.1 Introduction	168
9.1.1 Group and Subject-Specific Analysis of Functional Integration	169
9.1.2 Using a Multi-Subject Network to test Differences and Commonalities	170
9.1.3 The Neural Correlates of Single Word Reading	173
9.2 Methods	175
9.2.1 Specification of the Multi-Subject Network	176
9.2.2 SEM Analysis	178
9.3 Results	183
9.4 Discussion	186
9.4.1 Goodness of Fit of the Multi-Subject Network	187
9.4.2 The Effect of Word Type on Effective Connectivity Over Subjects	187
9.4.3 Effective Connectivity and Inter-Subject Variability	188
9.4.4 Further Applications of the Multi-Subject Network Analysis	189

Chapter 10 General Conclusions

10.1 Summary of the thesis	191
10.2 BOLD saturation, efficiency-mediated and neurophysiological effects	194
10.3 Comparing event-related and epoch models	196
10.4 Consistent and inconsistent effects across subjects	200
10.5 Future work	202
10.5.1 The coupling between neuronal activity and haemodynamics10.5.2 The effect of the task on the neurophysiological effects of rate and duration10.5.3 Alternative analytical techniques for the estimation of effective connectivity.	202 203 203

Appendix	
----------	--

eferences

List of Figures

Chapter 1

Figure 1.1 Fixed and distributed sampling of fMRI data	16
Figure 1.2 Rate effects in PET and fMRI from Rees et al. (1997)	27
Figure 1.3 Examples of epoch and event-related models	31

Chapter 2

Figure 2.1 Spin and precession of a single proton	0
Figure 2.2 Randomly orientated, parallel and anti-parallel spins	2
Figure 2.3 Effect of spin echo on phase coherence	4
Figure 2.4 The effect of a spin echo on phase coherence 4	17

Chapter 3

Figure 3.1 Data pre-processing	
Figure 3.2 Statistical parametric mapping	62
Figure 3.3 Model for modulatory effects in SEM	74

Chapter 4

Figure 4.1 Simulated BOLD and rCBF responses to a pair of words	.92
Figure 4.2 Effect of epoch length on rCBF and BOLD estimates	.93
Figure 4.3 Effect of SOA on rCBF and BOLD estimates	.95
Figure 4.4 Effect of stimulus amplitude on rCBF and BOLD estimates (for varying E ₀)	.97
Figure 4.5 Effect of the relative size of early and late neuronal components	.99

Chapter 5

Figure 5.1	Examples of	blocked and randomised sequences	111
------------	-------------	----------------------------------	-----

Chapter 6

Figure 6.1 Regions showing effects of reading and word rate	
Figure 6.2 Effect of stimulus amplitude on rCBF and BOLD estimates (for a f	ixed E ₀)138

Chapter 7

Figure 7.1 Different onset and offset for epoch and event-related models	141
Figure 7.2 Activations detected by the event-related model only	

Chapter 8

Figure 8.1 Dual route and connectionist models of reading	
Figure 8.2 Regions showing an effect of pseudowords relative to words	159

Chapter 9

Figure 9.1 Anatomical model of reading developed by Price (2000)	174
Figure 9.2 Subject-specific regions that entered into the SEM	177
Figure 9.3 Model for the modulatory effect of word type on connectivity	180
Figure 9.4 Multi-subject network	182
Figure 9.5 Graphical representation of the SEM results	186

Chapter 10

Figure 10.1 Relationship betw	veen different types of effects	
Figure 10.2 Comparison of ep	och and event-related models for di	ifferent SOAs199

List of Tables

Chapter 4	
Table 4.1 The six parameters of the haemodynamic model	89
Chapter 5	
Table 5.1 Positive effects of stimulus duration	118
Chapter 6	
Table 6.1 Effects of reading relative to fixation	
Table 6.2 Positive effects of stimulus rate	
Chapter 7	
Table 7.1 Effects explained by the event-related model only	146
Chapter 8	
Table 8.1 Previous studies on words vs. pseudoword reading	
Table 8.2 Effects of reading relative to fixation	157
Table 8.3 Effects of word type	159
Table 8.4 Subject-specific effects of words relative to pseudowords	160
Table 8.5 Positive and negative effects of stimulus duration	161
Chapter 9	
Table 9.1 The effect of word type on effective connectivity	

Chapter 1

Chapter 1

General Introduction

1.1 Introduction

This thesis investigates a number of factors that affect sensitivity to language activations in functional Magnetic Resonance Imaging (fMRI) measuring blood oxygenation level dependent (BOLD) contrast. Here I refer to sensitivity as the ability to detect experimentally-induced activations, which is a function of (i) the size of the effect of interest and (ii) the efficiency of the design with which the haemodynamic responses are estimated.

In the first part of this thesis, I investigate the impact of experimental design on the ability to detect language activations. Design parameters include stimulus rate, stimulus duration, stimulus size, epoch length and stimulus ordering. Crucially, these parameters may affect the sensitivity in multiple ways, including BOLD saturation, neurophysiological, and efficiency-mediated effects. I illustrate and discuss these effects by presenting biophysical simulations and fMRI studies of single word and pseudoword reading.

In the second part, I consider different types of analysis that can be used to estimate effects of interest from the data. First, I explore the relative sensitivity of eventrelated and epoch analyses in the context of blocked designs. Second, I address the notion that experimentally-induced effects may be detected not only as task-dependent changes in response amplitude but also as changes in connectivity amongst functionally

13

connected regions. These two complementary approaches are motivated by two fundamental principles of brain organisation: functional specialisation and functional integration. I present two studies in which I investigate the neural correlates of reading words and pseudowords in terms of functional specialisation and functional integration. Furthermore, in both studies I address the issue of inter-subject variability, which may be a critical determinant of sensitivity.

In this introductory chapter, I provide a brief overview of some of the main factors that affect the ability to detect activations in fMRI. This overview also serves as an introduction to the issues dealt with, in greater depth, in subsequent chapters. In the first part of the chapter, I discuss a number of issues related to data acquisition. The first condition required to detect activations reliably throughout the brain is unbiased sampling of the haemodynamic response, which I consider in section 1.2. Once unbiased sampling is achieved, the ability to detect activations will depend on the experimental design parameters used. These can affect sensitivity through BOLD saturation, neurophysiological, and efficiency-mediated effects, as discussed in section 1.3. The reasons why these effects may vary with imaging modality (i.e. Positron Emission Tomography vs. fMRI) are considered in section 1.4. In the second part of the chapter, I turn to issues in data analysis. In section 1.5, I discuss the use of event-related and epoch models in the context of blocked design fMRI. Finally, in section 1.6, I consider two complementary ways of analysing functional neuroimaging data, based on the principles of functional specialisation and functional integration.

1.2 Response Sampling

The first condition required to estimate activations reliably throughout the brain with fMRI is unbiased sampling of the haemodynamic response. FMRI acquires whole brain images by collating data from slices of the brain that have been sampled sequentially over time (rather than simultaneously as in Positron Emission Tomography). If the temporal relationship between data acquisition and stimulus presentation is fixed, different slices are acquired at fixed peristimulus times throughout the experiment. This may lead to a bias in the estimated activation, which depends on when the evoked BOLD response is sampled relative to the stimulus onset asynchrony (SOA). For instance, in event-related fMRI, sampling the peak of the evoked haemodynamic response will lead to an overestimate of activation whereas sampling any initial dip or the undershoot will lead to biased underestimates.

Biased sampling of peristimulus time was not originally considered to be a problem in blocked design fMRI because of the implicit assumption that steady-state dynamics are attained within each block. However, the evoked BOLD response may not conform to steady-state dynamics even in the context of blocked design fMRI. For instance, a steady-state response may be obtained when regions express an enduring temporal activation profile but not when the response is transient relative to the SOA. Here a periodic and dynamically modulated response may be obtained which is a function of peristimulus time. The estimated activation will therefore depend on when the response is sampled in the SOA. For instance, sampling the peaks will lead to an overestimate of steady-state activation whereas sampling the troughs will lead to biased underestimate (see upper panels of Figure 1.1).



Figure 1.1. Differences between fixed (upper panels) and distributed (lower panels) sampling when an SOA of 4 seconds is used, in the context of blocked design fMRI. In this simulation, the actual haemodynamic response (thin black line) is obtained by convolving an underlying stick function, which represents the stimulus onsets, with a synthetic haemodynamic response function. The thick gray line represents the estimated haemodynamic response. Upper panels: when acquisition and stimulus presentation are fixed, the estimated average activation is higher for slice 2 than slice 1 even though the actual haemodynamic response is the same. The bias is due to sampling the data slightly later in slice 2 than slice 1 in each scan. Lower panels: when the sampling is distributed over the SOA, a less biased estimate is obtained. From Price *et al.* (1999).

However, in both event-related and blocked design fMRI, an unbiased estimated of the response can be ensured by sampling the data in a distributed way over peristimulus time (see lower panels of Figure 1.1). Distributed sampling can be achieved in two ways. First, the presentation of the stimuli can be jittered by varying the SOA (Josephs and Henson, 1999; Josephs *et al.*, 1997). For example, the scan repetition time (i.e. the time required to acquire one image) might be 2 seconds and the SOA might be 2 seconds \pm 150 milliseconds. Alternatively, distributed sampling can be achieved by using a scan repetition time and a SOA that are not integer multiples of each other.

The importance of distributed sampling, even in the context of blocked design fMRI, has been illustrated empirically. Price *et al.* (1999), using a word rhyming paradigm, showed that activation in Wernicke's area was only detected when data acquisition was distributed throughout the peristimulus time. This suggests that a steady-state BOLD signal did not develop even if stimuli were only 3 seconds apart, an observation that is consistent with the finding that neuronal responses in Wernicke's area express transient temporal profiles (Cannestra *et al.*, 2000). Veltman *et al.* (in press) made direct statistical comparisons between data acquired from one or multiple time points and found that non distributed sampling led to biased estimates of activations in a number of prefrontal and temporal regions.

It should be noted that the risk of incorrect inference, due to biased estimates of activation, is greater in single-session experiments than in experiments in which independent sessions are pooled from the same subject or from different subjects. This is because variability in acquisition parameters (e.g., different positioning in the scanner) will introduce some variability in the peristimulus time at which data are

17

acquired. The mean response in a pooled analysis will therefore be a less biased estimate of the true response.

To summarize, when the temporal relationship between data acquisition and stimulus presentation is fixed, some slices are less sensitive to activations than others. However, by sampling the data in a distributed way over the SOA, one can ensure that all components of the BOLD signal are detected over scans. In the fMRI experiments presented in this thesis, the SOA and the scan acquisition time used were not integer multiples of each other. This ensured that every 15-90 milliseconds of peristimulus time was sampled equally.

1.3 The Effects of Experimental Design in fMRI

Once unbiased sampling of the haemodynamic response is ensured, the sensitivity with which activations are detected will depend on the experimental design. This can be characterised in terms of the design parameters, which may affect the ability to detect significant activations in multiple ways. Below, I consider the impact of design parameters on sensitivity through BOLD saturation effects. I then distinguish between two further ways in which design parameters may affect the sensitivity to neuronal activations. I will refer to these as neurophysiological and efficiency-mediated effects.

1.3.1 BOLD Saturation effects

There is good evidence for nonlinearity in the blood oxygenation level dependent (BOLD) response measured in fMRI (Miezin *et al.*, 2000; Friston *et al.*, 1998; Pollmann *et al.*, 1998). This nonlinearity can be thought of as a "saturation" or "refractoriness" effect whereby the response to a run of events is smaller than would be predicted by

the summation of responses to each event alone. This saturation is believed to arise in the mapping from cerebral blood flow to BOLD signal (Friston *et al.*, 2000; Huettel and McCarthy, 2000; Rees *et al.*, 1997), although may also have a neuronal component.

Since the degree of saturation in the evoked BOLD response depends on prior activation history, nonlinearity is clearly a function of the design parameters used. For instance, for a fixed number of stimuli, faster stimulus rates will increase the nonlinearity in the BOLD signal which in turn will decrease the sensitivity to experimentally-induced activations. Alternatively, for a fixed SOA of a few seconds or less, long epoch lengths will increase the nonlinearity in the BOLD signal which in turn will decrease sensitivity. In Chapter 4, I explore the impact of BOLD saturation on sensitivity across a range of different design parameters, by performing biophysical simulation studies based on a dynamical model of perfusion changes (Friston *et al.*, 2000). Furthermore, I address the notion that neuronal responses themselves may vary, in terms of the relative contributions of early and late neural components, and show that this may affect the degree of nonlinearity in the BOLD signal.

1.3.2 Neurophysiological effects

Neurophysiological or "direct" effects pertain to the influence that one experimental parameter exerts on neuronal response. These are the effects of interest in most functional neuroimaging experiments, in which the neuronal correlates of sensorimotor and cognitive functions are investigated. As shown by previous studies (e.g. see Price *et al.*, 1997; Price *et al.*, 1996a), the neurophysiological effects of the experimental design

parameters differ from region to region. For instance, during single word reading, bilateral occipital and left superior parietal areas activate throughout the time a word is presented thereby showing a linear increase in activity with stimulus duration (Price *et al.*, 1997). In contrast, right hemisphere temporal and inferior parietal areas show a monotonic decrease in activity with increased duration (Price *et al.*, 1997). Investigating the neurophysiological effects of experimental design parameters not only identifies the optimum design parameters for maximising sensitivity in regions of interest, but also enables the segregation of brain regions showing differential responses. Examples of neurophysiological effects of stimulus ordering, rate and duration during single word and pseudoword reading are presented in Chapters 5, 6 and 8 respectively.

1.3.3 Efficiency-mediated effects

Efficiency-mediated or "indirect" effects pertain to the efficiency with which haemodynamic responses are estimated. Efficiency of response estimation is a measure of how reliable the parameter estimates are and depends on the design variance (i.e. a function of the contrast tested and the design matrix) and the error variance (the variance in the data which is not modelled by explanatory variables in the design matrix) (see Chapter 5 and Appendix for details). The relative efficiency of different experimental designs may be expressed as differences in standard error that can have substantial effects on the ensuing statistics. For instance, under the assumption that the error variance is independent of changes in the experimental design, blocked designs are more efficient than randomised designs (Friston *et al.*, 1999). This results in the standard error being smaller for contrasts testing for activations in blocked than

randomised designs. The *T* values will therefore be higher in the blocked designs even if the amplitude of the haemodynamic responses and the parameter estimates are identical. In a design that includes both blocked and randomised components, this means it should be possible to show significant effects within the blocked but not the randomised component and yet no interaction between the effects and presentation mode. This is illustrated in Chapter 5 of this thesis, in which an event-related study of single word reading is presented which involved acquiring data using an experimental design that embodied both blocked and randomised trials.

As stated above, efficiency can be factorised into design variance and error variance. It should be noted that the design variance (i) depends on the contrast and design matrix only, (ii) can be computed *a priori*, and (iii) is the same across the whole brain. In contrast, the error variance (i) depends on cognitive/physiological effects (e.g. the haemodynamic responses may be more variable in one context relative to another), (ii) can only be estimated by performing a statistical analysis, and (iii) is voxel-specific. This means that efficiency-mediated effects (i.e. the relative efficiency for two or more experimental designs) can only be quantified *a priori* by assuming that the error variance is independent of changes in the experimental design (Friston *et al.*, 1999). For this assumption to be met, the responses need to conform to a linear convolution model, which embodies two further assumptions: BOLD nonlinearities can be discounted and the form of the haemodynamic response function is the same for different experimental designs. When these assumptions are violated, differences in error variance may arise, thereby compounding the relative efficiency of the designs. In Chapter 5, a study is presented which investigated whether the relative efficiency of blocked and randomised

designs can be predicted under the assumptions adopted generally in the context of *a priori* estimation of efficiency. Differences in the error variance would indicate that both the error variance and the design need to be considered when estimating efficiency. Alternatively, if differences in error variance are negligible, the relative efficiency can be predicted from the design alone.

Critically, the choice of experimental parameters may affect sensitivity through both neurophysiological and efficiency-mediated effects. The contribution of efficiencymediated effects to the statistics means that one should never *compare statistics* to make inferences about differential responses. Rather, one should always make a *statistical comparison* of the data, generally testing for an interaction.

1.4 The Effects of Experimental Design in PET and fMRI

In the previous section, the ways in which experimental design parameters may affect sensitivity in fMRI were discussed. In this section, I discuss the differential ways that experimental design parameters can affect sensitivity in PET and fMRI. This is an important issue because the replication of results across imaging modalities identifies effects which have a face validity (i.e. true neurophysiological effects). The discrepancy between the effects of experimental design parameters in PET and fMRI has already been shown for auditory words. For instance, within the range of 10-90 words per minute (wpm), the regional cerebral blood flow (rCBF) in primary auditory cortex is proportional to auditory word presentation rate (Rees *et al.*, 1997; Price *et al.*, 1992), whereas the BOLD signal in the same region shows a saturable effect for high

presentation rates (Rees *et al.*, 1997). On the other hand, rCBF in the left posterior superior temporal gyrus (Wernicke's area) increases in response to auditory words irrespective of their presentation rate (Price *et al.*, 1992), whereas the BOLD signal in the same region is proportional to presentation rate (Dhankhar *et al.*, 1997).

The discrepancy between the dependence of PET and fMRI signals on stimulus rate may be due to (i) PET and fMRI measuring different physiological parameters (rCBF and deoxyhaemoglobin concentration) that do not share a common stimulus rate dependency, (ii) differences in signal acquisition, and (iii) differences in the nature of the baseline in PET and fMRI. These will now be discussed in turn.

1.4.1 Physiological Parameters measured in PET and fMRI

Understanding the nature of activity-dependent haemodynamic changes may help interpret the discrepancy between PET and fMRI responses when stimulus rate or other experimental design parameters vary. The presentation of a stimulus involves, in the first instance, neural activity in specific brain regions. This activity elicits (i) electrical signals in nerves to arterioles and (ii) the synthesis of Nitric Oxide. Nitric Oxide is an informational, as opposed to energetic, substance that mediates vasodilatation by rapidly diffusing from the site of neural activity to sites of action in the microvessels (see Friston, 1995a for a review; Iadecola *et al.*, 1993; Gally *et al.*, 1990). Both neurogenic and passive diffusion signals ensure muscle relaxation around vessels and mediate the coupling between neural activation and blood flow changes within a few hundred milliseconds. A modest increase in the cerebral metabolic rate of oxygen consumption is accompanied by a much larger increase in local blood flow. Because of this imbalance, the local deoxyhaemoglobin concentration decreases during brain activation. The increase in blood flow is the basis for measuring activations with PET, whereas the decrease in local deoxyhaemoglobin concentration is the basis for measuring responses with fMRI. When a second stimulus is presented, the coupling between neuronal dynamics and blood flow is believed to be unaffected by the occurrence of the first stimulus. In contrast, the coupling between neuronal activity and the BOLD signal depends on prior stimulation (Huettel and McCarthy 2000; Rees *et al.*, 1997). Buxton and colleagues (1998) have proposed that the presentation of a stimulus inflates a venous "balloon". Once this "balloon" is inflated, the rate at which diluted deoxyhaemoglobin is expelled in response to the underlying neuronal activity will be compromised and the BOLD signal will be reduced. The response to the second stimulus will therefore be different in PET and fMRI because the physiological parameters they measure, rCBF and deoxyhaemoglobin concentration, do not share a common dependency on the underlying neuronal activity.

The reduced BOLD responsiveness to a second stimulus, or BOLD refractoriness, may account for the saturable effect of auditory stimulus rate detected in fMRI (Rees *et al.*, 1997) but not in PET (Rees *et al.*, 1997; Price *et al.*, 1992). This dependency of fMRI sensitivity on previous stimulation, and therefore on stimulus rate, makes the choice of the appropriate SOA a crucial factor in designing fMRI experiments (Friston *et al.*, 1999). In Chapter 4, I introduce a new class of nonlinearity that distinguishes BOLD from rCBF signals, namely haemodynamic "rectification". This refers to the possibility that BOLD signals are less sensitive to rCBF decreases than they are to increases. This would be expressed, in the context of parametric designs, by a

differential sensitivity to the positive and negative effects of rate between fMRI and PET. This is exactly what I observed in the data reported in Chapter 6.

1.4.2 Data Acquisition in PET and fMRI

If we ignore true differences in rate-dependence between PET and fMRI due to nonlinearities inherent in the latter, then PET and fMRI should show the same ratedependence. However, this does not imply that a given dependency will be detected with equal sensitivity by PET and fMRI (i.e. although the regression slopes of signal dependence on rate may be identical, the significance of T values may not be). Differences in sensitivity between PET and fMRI may reflect differences in the acquisition of the two data sets. In PET haemodynamic responses are measured continuously and summed over a 60-90 second scan (an epoch of stimuli) at a given rate. Statistical inferences depend on the mean signal per epoch and the between epoch variability, and are not affected by the distribution of the haemodynamic responses within the epoch. On the other hand, fMRI data are acquired by sampling the haemodynamic response every 2 to 4 seconds. This means that data are acquired 20-40 times more often than in PET over the same time. In this case the statistical inference is affected by the distribution of the haemodynamic responses within the presentation epoch. In short, fMRI inferences are based on within-epoch error variance, which is used to assess the variability in between-epoch estimates and their significance. In PET, inference is based directly on between-epoch error variance. If the within-epoch error is substantially smaller than between-epoch error, fMRI inferences may be, apparently, much more sensitive. This difference in sensitivity may have contributed to the

proportional effect of stimulus rate detected in Wernicke's area using fMRI (Dhankhar *et al.*, 1997) but not in PET (Price *et al.*, 1992) during auditory word processing (but see (c) below).

1.4.3 The Nature of the Baseline in PET and fMRI

Finally, differences between PET and fMRI may be due to differences concerning the behavioural baseline. During PET data acquisition, each scan is clearly separated from the others, the inter-scan interval being 8 minutes or more. Subjects are therefore aware that stimuli will not occur while the resting scans are acquired. During fMRI data acquisition, however, resting time and stimulus presentation usually alternate without long intervals of rest. While the resting scans are acquired, subjects are therefore aware that stimuli may be presented at any time and their attention or expectation may enhance activity in stimulus processing areas thereby reducing activation differences in fMRI relative to PET (Veltman *et al.*, 2000; Chertkow *et al.*, 1993). If 0 wpm (the baseline of 0 words per minute) is included in the regression, the difference between PET and fMRI baselines may introduce or exaggerate nonlinearities in rate effects in PET relative to fMRI. This is well illustrated in a study by Rees *et al.* (1997) (see Figure 1.2).



Figure 1.2 The effect of stimulus rate on rCBF (upper panels) and BOLD (lower panels) in the left and right auditory cortex during single auditory word processing (as reported by Rees *et al.*, 1997). It can be seen that the difference between PET and fMRI baselines introduced or exaggerated nonlinearities in rate effects in PET relative to fMRI if rest is treated as 0 *wpm*.

Baseline differences may explain why in Wernicke's area, the effect of increasing the rate of auditory word presentation was found to be: (i) proportional to rCBF in Mummery *et al.*'s (1999) study, when subjects expected to hear words in the baseline (although none arrived); (ii) proportional to the BOLD response in Dhankhar *et al.*'s (1997) study when the baseline was rest in fMRI and subjects expect stimuli to be delivered within seconds; and (iii) non-proportional to rCBF in Price *et al.*'s (1992) study when the baseline was rest with no expectation that any stimuli would be presented. In Chapter 6, I present a study which investigated whether the differences between PET and fMRI rate dependency observed in the auditory domain are also seen during visual word processing. Because of the potential effect of baseline on estimating proportional or non-proportional rate effects, 0 wpm (the baseline) was deliberately excluded from the analysis of rate effects in the present experiment.

In summary, (i) PET and fMRI measure different physiological parameters which may explain why the effect of stimulus rate in sensory cortex is nonlinear with BOLD and linear with rCBF; (ii) different variance component estimation (within-epoch *vs.* between-epoch) may explain why the effect of stimulus rate can be declared significant in BOLD and not in PET; and (iii) differences in attentional set between PET and fMRI baselines may account for further discrepancies if rest is entered into the regression parametrically. It should be noted that these factors correspond to BOLD saturation, efficiency-mediated and neurophysiological effects as discussed in the previous section.

1.5. Epoch vs. Event-Related Analysis

While the first part of this chapter considered a number of issues related to data acquisition, the last two sections focus on data analysis. As mentioned in section 1.2 of this chapter, a distinction is made in fMRI between event-related designs, in which stimuli of different types are intermixed, and blocked designs, in which stimuli of the same type are presented in blocks. Effects of interest in blocked designs are usually modelled with some form of boxcar regressor convolved with a synthetic haemodynamic response function (HRF). Implicit in this model is the assumption that steady-state synaptic activity and haemodynamics are attained within each block. In contrast, effects of interest in event-related designs are modelled by convolving each trial onset (i.e. a stick function) with a synthetic HRF or temporal basis set. Here the haemodynamic response to stimulus-induced neuronal transients is modelled without assuming constant within-block activity.

In this section I discuss the use of epoch and event-related analyses of data collected in blocked design fMRI and note that the two statistical models may provide differential sensitivity to experimentally-induced effects. This is because the two types of models may differ with respect to the temporal shape of the predicted response in two fundamental ways. First, for relatively long SOAs (e.g. 3 seconds or more), an epoch model will predict steady-state dynamics whereas an event-related model will predict a periodic and dynamically modulated response, as represented in Figure 1.3c. If the measured BOLD response to a train of stimuli is too transient relative to the SOA, the response to any stimulus will have died away before the presentation of a subsequent stimulus. This will result in a periodic and dynamically modulated response as predicted

by the event-related model, which would explain the data better than the epoch model. However, if the measured BOLD response has an unusual (unpredicted) shape, the epoch model is less likely to be out of phase and may explain the data better than the event-related model. It should be noted that the two types of models will differ, as represented in Figure 1.3c, only when the SOA is long relative to the scan acquisition time.

Second, event-related and epoch models may differ with respect to the onset and offset of the predicted response to a train of stimuli. Figure 1.3 illustrates differences in onset and offset latencies between epoch (red solid line) and event-related (blue broken line) models, when stimulus onset asynchrony (SOA) is 3 seconds. It can be seen that, although the two models assume the same amount of integrated synaptic activity, the event-related model expresses higher frequencies than the epoch model (Figure 1.3a). Once the assumed synaptic activity is convolved with the HRF (Figure 1.3b), the two models predict differential haemodynamic responses, with the epoch model reaching its peak later and returning to baseline sooner than the event-related model (Figure 1.3c). This corresponds to differential response onsets and offsets for the event-related and the epoch model. When the epoch model is orthogonalised with respect to the event-related model (Figure 1.3d), effects which are modelled by the event-related but not by the epoch model are identified. It appears that the event-related model explains changes in activity at the beginning and at the end of the block that are not accounted for by the epoch model.

Chapter 1



Figure 1.3 Difference in onset and offset latencies between epoch (red solid line) and event-related (blue broken line) models, when the SOA is 3 seconds. (a) Neuronal models for epoch and event-related models. (b) The neuronal models are convolved with the HRF to predict the haemodynamic responses. (c) Differential haemodynamic responses predicted by the epoch and the event-related model. Although the areas under curves are the same, the epoch model reaches its peak later and returns to baseline sooner than the event-related model. (d) When the epoch model is orthogonalised with respect to the event-related model, effects which are modelled by the event-related model by the epoch model are identified. It can be seen that the event-related model explains changes in activity at the beginning and at the end of the block that are not accounted for by the epoch model.

In short, although both event-related and blocked models can be used to analyse data acquired using a blocked design paradigm, the two types of models may differ because (i) the epoch model always predict steady-state dynamics whereas an event-related model may predict a periodic and dynamically modulated response; and (ii) the event-related model may reach its peak earlier and return to baseline later than the epoch model. It should be noted that the differences between event-related and epoch models are a function of stimulus rate. These differences would disappear in the limit of very fast presentation rates because the epoch and event-related regressors would be the same.

A critical question is which model one should use to best explain the data, thereby maximising the sensitivity with which statistical inferences are drawn. In Chapter 7 of this thesis, I explore the impact on sensitivity of differences in response onsets and offsets between event-related and epoch models, using a blocked design reading paradigm. I show that the differential onset of the predicted BOLD response in the event-related and epoch model may affect the ability to detect activations in critical language areas.

1.6 Characterising Experimentally-Induced Effects as Neuronal Interactions

Here I address the notion that experimentally-induced effects may be detected not only as task-dependent changes in response amplitude but also as changes in coupling amongst functionally connected regions. These two complementary types of analysis are motivated by the two well established principles of brain organisation, functional specialisation and functional integration.

1.6.1 Functional Specialisation and Functional Integration

In the 19th century, investigations of brain functions were dominated by the idea that different regions serve different functions, a principle known as functional specialisation. This approach was driven largely by the data available to scientists of that era, who studied neurological patients with circumscribed lesions impaired in one or more specific abilities. In the first part of the 20th century, the idea of functional specialisation was overtaken by the doctrine of "mass action", which proposed that higher abilities depended on the functioning of the brain as a whole (Lashley, 1929). However, the techniques available at the time did not allow scientists to make any progress studying the brain as a whole. By the end of the 20th century, the principle of functional specialisation supervened, especially in the domain of cognitive neuropsychology (Shallice, 1988).

The first functional neuroimaging experiments of cognition were based on the principle of functional specialisation (Petersen *et al.*, 1990; Posner *et al.*, 1988; Petersen *et al.*, 1988). These studies investigated the neural correlates of written word processing by assuming that reading goes through a single series of discrete and independent stages. By comparing responses to a hierarchy of stimuli (e.g. false fonts vs. letters; letter strings vs. words, etc.), each stage was isolated and the associated brain region identified. The characterisation of brain activity in terms of functional specialisation is still predominant in functional imaging studies of sensorimotor and higher cognitive functions. A critical point is that this approach assumes that different areas communicate with each other. For instance, the neural correlates of semantic processing can be identified using written words by virtue of the assumption that visual regions interact

with semantic regions. However, characterising brain activity in terms of functional specialisation does not reveal anything about how different brain regions communicate with each other. In other words, functional specialisation can only provide a limited account of the neuronal substrate of the processes investigated.

1.6.2 Functional and Effective Connectivity

Alternative approaches have therefore been developed in the context of functional neuroimaging to investigate the integration of functionally specialised areas. These approaches use a number of concepts and definitions which are derived from multiunit micro-electrode recording of separable spike trains (Aertsen and Preissl, 1991; Gerstein and Perkel, 1969). For instance, functional connectivity is defined as the "temporal correlations between spatially remote neurophysiological events" (Friston *et al.*, 1993a), whereas effective connectivity is defined as the influence one neural system exerts over another either directly or indirectly (Friston *et al.*, 1993b). Functional connectivity is a statement about the observed correlations and does not provide any insight into how these correlations are mediated. At the level of multiunit micro-electrode recordings, functional connectivity can result from stimulus-locked transients, evoked by a common afferent input, or may reflect stimulus-induced phasic coupling of neuronal assemblies, mediated by synaptic connections. In contrast, effective connectivity is closer to the notion of connection, either at a synaptic or cortical level.

Investigating the functional and effective connectivity within a network may provide critical information on the neuronal correlates of the processes investigated. For instance, across different tasks a number of brain areas may show the same activity

34

pattern but serve different functions because the functional integration amongst them is different. In other words, functional and effective connectivity analyses may be *sensitive* to neuronal interactions that are important for the task performance but cannot be detected using a "cognitive subtraction" approach (for empirical examples, see Della-Maggiore *et al.*, 2000; McIntosh, 1999).

1.6.3 Inferring Effective Connectivity

The distinction between functional and effective connectivity speaks to the fact that temporal correlations between neurophysiological events in separate neural systems (i.e. functional connectivity) may or may not be due to the influence that one neural system exerts over another (i.e. effective connectivity). For instance, if two regions A and B receive common inputs from a third area C, then there will be covariance between A and B without an influence of A on B and *vice versa*. These common inputs may be either neuromodulatory or task specific neural inputs from other cortical areas. However, distinguishing unambiguously between functional and effective connectivity may not be possible simply by looking at the data. Below I consider two possible ways of inferring causality (i.e. effective connectivity) and the relative limitations.

A first way of inferring causality may involve considering temporal precedence. However, in the context of fMRI, the influence of one region over another may lead to different temporal responses at the neuronal level but may not result in temporal precedence at the haemodynamic level, because of the temporal smoothing embedded in the coupling between synaptic activity and perfusion changes. In extreme cases, one region may cause activity in another and yet, because of different latencies in the

35
coupling between synaptic activity and perfusion changes, the haemodynamic response in the "target" region may occur before the haemodynamic response in the "source" region. Even in the context of EEG and MEG studies, which provide a temporal resolution in the order of milliseconds, the distinction between functional and effective connectivity may not be possible since temporal precedence is not unambiguously linked to causality.

A second way of inferring causality may involve "perturbing the system", for instance using transcranial magnetic resonance imaging (TMS). TMS uses an externally-generated changing magnetic field, applied by a coil placed over the subject's head, to induce electric currents which can be used to disrupt focal neuronal activity in a region of interest (Barker *et al.*, 1985). This will affect activity not only in this region, but also in a number of connected areas depending on the underlying effective connectivity. However, it may be difficult to define precisely the spatial extent of TMS (Walsh and Rushworth, 1999). Furthermore, the way in which TMS affects neuronal function is still not well understood.

These difficulties are the reason why, in practice, the characterisation of a data set in terms of either functional or effective connectivity usually depends on the assumptions and the hypotheses that motivate the study. The characterisation of brain activity in terms of functional connectivity is "model free" whereas characterising brain activity in terms of effective connectivity requires a "causal model" in which regions and connections of interest are specified by the researcher (Horwitz *et al.*, 2000; Friston, 1995b). These causal models are usually derived by combining neuroanatomical, neuropsychological and functional neuroimaging data.

In short, although functional imaging experiments are generally based on the principle of functional specialisation, this type of analysis does not provide any information about how different brain regions communicate with each other. In contrast, the analysis of functional and effective connectivity is sensitive to neuronal interactions and may help explain behavioural data. In Chapters 8 and 9, I investigate the effect of word type (i.e. words vs. pseudowords) on brain activity by analysing the data using both cognitive subtraction and effective connectivity approaches. Furthermore, in both studies I address the issue of inter-subject variability, which constitutes an important error variance component in random effect analyses.

1.7 Summary

In this chapter, I have reviewed some of the main factors that affect sensitivity in functional Magnetic Resonance Imaging. In the first part of the chapter, I have discussed a number of issues related to data acquisition. First, multi-slice data acquisition may lead to an overestimate or underestimate of activation when only the peak or the undershoot is sampled. In order to avoid biased estimation, the experiments presented in this thesis were performed using distributed sampling. Second, the sensitivity with which effects are detected will depend on the experimental design. Chapters 4,5,6 and 8 of this thesis investigate the multiple ways in which experimental design parameters may affect the ability to detect language activations.

In the second part of the chapter, I have focused on data analysis. I have addressed the notion that both event-related and epoch models can be used in blocked design fMRI. However, the two types of models may provide differential sensitivity to experimentally-induced effects, as I show in Chapter 7. Finally, I have considered the analysis of functional imaging data based on the principles of functional specialisation and integration. Examples of these approaches are presented in Chapters 8 and 9, in which the effect of word type (i.e. words vs. pseudowords) during reading is investigated in terms of changes in response amplitude and effective connectivity.

Chapter 2

Chapter 2

Methods: Functional Magnetic Resonance Imaging

2.1 Introduction

The experiments reported in this thesis were conducted using magnetic resonance imaging (MRI). This chapter describes the principles of MRI data acquisition. MRI is part of a larger family of techniques used to characterise brain structure and activity (Orrison *et al.*, 1995; Toga and Mazziotta, 1995), that also includes event-related elecrophysiological potentials (ERP), magnetoncephalography (MEG), magnetic resonance spectroscopy (MRS), positron emission tomography (PET) and single photon emission computed tomography (SPECT).

2.2 The physics of NMR

Functional MRI is a special form of the magnetic resonance imaging (MRI) technique that is used widely in clinical practice. Both fMRI and MRI rely on the principle of nuclear magnetic resonance (NMR), discovered independently by Bloch and Purcell in the 1940s (Bloch *et al.*, 1946; Purcell *et al.*, 1945). The first application of NMR as a topographic imaging technique was by Lauterbur (1973), who applied NMR of hydrogen atoms to histological specimens. This early study provided a basis for the use of MRI to image the structure of human tissue and, later, for the use of fMRI to measure brain activity. The basic physical principles underlying MRI and fMRI are very similar.

2.2.1 Spin

The nucleus of an atom contains protons - particles that have a positive electrical charge. Protons possess the quantum quality spin, i.e. they are constantly turning or spinning around an axis. The positive electrical charge, being attached to the proton, spins around inducing a magnetic field with north-south polarity along the spin axis (i.e. magnetic vector). In the absence of an external magnetic field, the individual spins are randomly orientated and the bulk material has no net magnetisation. However, in an externally applied magnetic field, B₀, the individual spins align with the external magnetic field like a compass needle in the magnetic field of the earth. The spins are not completely aligned with the direction of the magnetic field B₀. Rather the protons revolve, or precess, around the field direction (Figure 2.1).



Figure 2.1 Spin and precession of a single proton. The proton possesses the quantum quality spin (red arrow). In an external magnetic field, the proton precesses (black curved arrow) around the longitudinal magnetisation vector (B_0) .

The precessional frequency, ω_0 , is governed by the strength of the magnetic field B_0 , which is given in Tesla (T), and a gyromagnetic constant for each type of nucleus γ , where:

$$\omega_0 = \gamma B_0$$

This frequency is also known as the Larmor frequency. MRI techniques measure the effects of changing the spin of particular atomic nuclei, which have an odd number of protons. In living organisms, the most abundant source of protons comes from hydrogen atoms in the form of water. For hydrogen in water molecules, the gyromagnetic constant γ is approximately 40 MHz/T, which results in a precessional frequency (ω_0) of 80 MHz at B₀ ~ 2 Tesla. This means that, when an external magnetic field of 2 Tesla is applied, the protons precess around the magnetic field direction 85 million times per second.

A spinning proton can have different energies depending on its orientation with respect to the applied magnetic field B_0 . For the simple spin system of hydrogen, the spinning nucleus can have two orientations with respect the applied magnetic field B_0 , either parallel to it or anti-parallel to it (see Figure 2.2). These two types of alignment are on different energy levels. A spinning nucleus which aligns with the direction of the external field is associated with a low, ground energy state. A spinning nucleus which points exactly in the complete opposite direction is associated with a high energy state. The quantum of energy required to flip orientations (e.g. from parallel to anti-parallel) depends on the field strength. The generation of the signals in MRI is based on differences between these energy states.

Chapter 2



Figure 2.2 In the absence of an external magnetic field, the individual spins are randomly orientated (left). However, in an externally applied magnetic field B_0 , the individual spins align parallel or anti-parallel to the magnetic field B_0 (right). The spin axis of a nucleus (broken line) is not identical to B_0 , as the nucleus will precess around B_0 .

2.2.2 Net magnetisation

The sum over all the nuclei in an object volume gives the net magnetisation for the object. The description of the net magnetisation is based on a co-ordinate system, with the *z* axis being in the direction of the applied magnetic field B_0 . The resting net magnetisation is called the equilibrium magnetisation. In equilibrium, more spins are in the ground than in the high energy state. Summing the contributions of individual magnetic vectors will, therefore, give a net magnetic vector M_0 along the direction of the applied magnetic field (Figure 2.3 (left)). In addition, a small part of the rotating magnetisation of each nucleus has a component projecting into the *xy* plane (i.e. transverse to the field direction). However, there is no resting net 'transversal magnetisation' because opposing magnetic forces in the transverse plane cancel to zero.

2.2.3 Radiofrequency magnetic fields

Spins can be excited from the ground to the high energy state by applying an oscillating radiofrequency electromagnetic field (B_1) perpendicular to the main magnetic field (B_0 , applied in the *z* plane). To achieve the most efficient transfer of energy, the oscillation frequency of the B_1 field should be the same as the spin (Larmor) frequency. This will result in a decrease of the longitudinal field M_0 (see Figure 2.3 (right)). The transfer of energy from the ground to the high energy state, which occurs when the oscillating radiofrequency B_1 and the protons have the same frequency, is called "resonance".

A second effect of the radiofrequency (RF) pulse B_1 is that it rephases the protons so that they are brought into coherence (in phase). Individual magnetic vectors in the transverse plane no longer cancel to zero, but point at the same direction at the same time. This will result in a new magnetic vector M_1 pointing in a transverse direction into the transverse plane (Figure 2.3 (right)). By varying the amplitude and duration of the RF exposure, typically delivered in millisecond pulses, any desired angle can be produced. For most angles from the equilibrium longitudinal magnetisation there will be a non-zero component of the transversal magnetisation in the *xy* plane. For example, with enough energy, the magnetic vector can be rotated in the *xy* plane. This is referred to as a 90° pulse, producing a 90° rotation of the net magnetisation – a 90° angle. It is the transverse component of the new magnetic vector that gives rise to a detectable NMR signal. Following the 90° pulse, the magnetisation vector will rotate about B_0 in the *xy* plane with the spin resonance frequency as represented in Figure 2.3 (right). This is observable because the oscillating magnetisation vector induces an electric current in a coil positioned in the *xy* plane (Faraday's Law). MRI systems are designed to measure the transversal magnetisation, so the receiver coils, which may be the same as those used to apply the RF pulses, are sensitive to the transverse component only. The amplitude of the detected signal is dependent on the number of protons in the sample (the proton density). The greater the proton density, the greater the magnetisation hence the greater the signal detected by the RF coils.



Figure 2.3 In a strong magnetic field B_0 along the z axis, a new magnetic vector M_0 along the external field is established (left). Sending in an RF pulse B_1 causes a new transversal magnetisation M_1 while magnetisation along the external field decreases (right). Red arrows indicate the axis of spin of individual nuclei that originate the net magnetisation for the object.

2.2.4 Relaxation

After excitation, the spin system will ultimately give off the absorbed energy and return to its equilibrium state with the magnetisation vector returning to its original position along the z axis. Hence, the strength of the signal detected by the RF coils, which is directly related to the amount of net transversal magnetisation, will gradually decay to zero. The molecular environment of the nuclei is reflected in the time variation of signal amplitude as the nuclei return to equilibrium. This release of energy happens gradually and may occur in two ways: (i) energy can be given up to neighbouring molecules in the surrounding environment (spin-lattice relaxation), and (ii) energy can be given up to nearby nuclei (spin-spin relaxation).

T1 relaxation: Spin-lattice or T1 relaxation describes the regrowth of the magnetisation vector along the z axis. This regrowth is an exponential process described by the T1 time constant. Protons that have been excited to the higher energy state cause regrowth of the magnetisation vector along the z axis by dissipating energy to molecules of the surrounding structure ('lattice') as heat. The exact composition of the environment will affect T1. For example, the protons in water have a longer T1 than those in fat because the carbon bonds in fat resonate near the Larmor frequency, which facilitates the transfer of energy to the lattice. In the human brain, the different water content of grey and white matter (71% and 84%, respectively) means that T1 contrast can be used to provide contrast between these two tissues. T1 also depends on the strength of the magnetic field: the stronger the magnetic field, the faster the frequency with which protons precess. The frequency of

precession is directly proportional to the time required for the protons to realise their energy (i.e. T1).

T2 relaxation: Spin-spin or T2 relaxation describes the disappearance of coherence of the magnetic vector in the *xy* plane (the transversal magnetisation, M_1), occurring at a different rate to the recovery of magnetisation along the *z* axis. The term spin-spin refers to the fact that interactions between protons determine the rate of T2 relaxation. No energy is actually lost; rather, energy is exchanged between protons, and there is a loss of "order" or entropy. As neighbouring spins pass energy from one to another, their rotations become desynchronised. Their slightly different rotation frequencies result in build-up of phase differences that gradually decrease M_1 .

T2* relaxation: As described below, a non-uniform magnetic field B_1 (i.e. a magnetic field which has a different strength at each point of the object's cross section) is applied to produce tomographic MR images. However, using a non-uniform magnetic field will decrease the net transversal magnetisation M_1 . This is because spins in different parts of the object will rotate at different frequencies in a non-uniform field, thereby cancelling each other out. This loss of transversal magnetisation due to inhomogeneous fields is characterised by the exponential time constant T2'. The value of T2' is determined by the non-uniform magnetic field applied and further field inhomogeneitis caused by the properties of the object itself. Generally T2' is shorter than T2, i.e. signal decay due to inhomogeneous fields is faster than that due to spin-spin relaxation. T2* relaxation (as measured in fMRI) reflects the combination of T2 and T2' signal decays.

The magnetic field contribution to the inhomogeneity giving rise to the T2' signal decay can be refocused with a 'spin echo', a 180° pulse along an axis in the *xy* plan at some time *t* after the initial 90° excitation pulse. When a spin echo pulse is used, protons turn around so that they precess towards their original phase or orientation. This increases phase coherence which results in greater transversal magnetisation. At the time 2*t* (also called TE), all spins have the same phase and the M₁ is at its maximum again (see Figure 2.4). Although the 180° pulses cancel out T2' effects due to magnetic field inhomogeneities, the recovered T2* signal will be still less than its original height as determined by the proton density. This is due to T2 relaxation, the second contribution to M₁ magnetisation dephasing, which cannot be refocused with a spin-echo because it results from spatial and temporal variations in the intrinsic magnetic environment of each spin. In other words, signal will decay during TE with a time constant T2.



Figure 2.4 After the RF pulse is switched off, the protons dephase (left). The 180° pulse causes them to precess towards their original orientation and so they rephase again (right).

Echoes can also be formed without 180° RF pulses by using magnetic field gradients to dephase and then rephase the spins. Adjusting the time of the dephase/rephase gradient balance allows the time of echo formation to be changed. This type of 'gradient echo' is often used in fast imaging sequences where it can produce faster echoes and use less RF power than 180° echos. All T2* data presented in this thesis were acquired using this gradient echo technique.

2.2.5 Image formation: frequency and phase encoding

Using a homogenous field B_0 will not produce tomographic MR images of a sample. This is because all protons will experience (roughly) the same magnetic field and, hence, the frequencies of their emitted signal will all be identical. In MRI, a nonuniform magnetic field B_1 is applied so that the resonance frequencies of spins vary within the sample. In other words, as the gradient field varies across the sample, protons will emit different (Larmor) frequency signals according to their spatial position. For each frequency component of the measured signal, the known value of the applied gradient strength and direction can be used to calculate the position from which the signal came. The amplitude of a frequency in the spectrum of observable resonance frequencies is an index of the number of spins resonating at that particular frequency.

Combining a non-uniform magnetic field with an echo pulse of the appropriate frequency and bandwidth can excite a small slice of the sample, allowing a slice by slice investigation. However, if the signal was acquired by measuring frequencies only, its spectrum would be a one-dimensional projection of spin density for the spins in the selected slice. To produce a two-dimensional image of spin densities in the selected slice, encoding along a second axis is required. While along the first axis locations are encoded by frequencies, along the second axis they are encoded by phase. Locationdependent phase is achieved by temporarily switching on a further (linear) gradient along the second axis. An effect of any gradient is an enhanced dephasing of transversal magnetisation (M₁). During gradient application, local magnetisation vectors will rotate with different frequencies depending on their positions within the gradient. These spins will also express phase differences among their magnetisation vectors dependent on their positions along this second axis. The duration of the applied phase encoding gradient dictates the degree to which local transversal magnetisations are dephased. A series of increasing gradient pulse lengths will enable a reconstruction of the frequencies giving rise to the dephasing of transversal magnetisation. Hence, despite the fact that phase is being manipulated in the second axis, the amplitudes of spin frequencies are again determined and expressed as a spin density projection along the phase encode axis.

In summary, a typical imaging procedure starts with slice-selective excitation by the temporary application of a slice-selection gradient. Frequency encoding is determined by applying the frequency encoding gradient in the x axis (conventionally) during acquisition. This is followed by phase encoding along a second orthogonal axis (y axis). During image acquisition, locations are encoded by both frequency and phase of the detected signal (i.e. the transverse component of the magnetic vector).

49

Chapter 2

2.2.6 Voxels

Step-wise increases in both gradients divide the sample into small cubes, or voxels (volume-elements). Spins in one voxel experience the same frequency and phase encoding. The signal of a given voxel is the sum of all spin contributions hence spins within a voxel cannot be distinguished from each other. The resolution of the image depends on the size of the voxels, which is determined by the step size of the gradients. Increasing the size of the voxel increases its signal and therefore its signal to noise. However, larger voxels are more likely to encompass groups of spins with very different behaviour, which could evoke a misleading signal (referred to as the partial volume effect).

2.2.7 Image contrast

Image contrast is based on the difference in signal intensity between areas of different structure or composition in an image. The MR signal intensity from a given voxel arises from a complex interaction of many different factors including T1 and T2 relaxation times, proton density, RF pulse characteristics and magnetic susceptibility (magnetic susceptibility refers to the fact that the net field experienced by a given nucleus depends on other magnetic spins or electron clouds in their environments). The relative contribution of some of these factors to the transversal magnetisation may be manipulated by controlling the timing of the RF pulses (known as pulse sequence parameters). The timing parameters are the repetition time (TR; the time required to acquire each image volume, i.e. the time between two consecutive 90° RF pulses) and echo time (see earlier). For example, a short TR and TE will emphasise the T1

characteristics of the tissue and produce a "T1 weighted" image. A long TR and TE will emphasise the T2 characteristics of the tissue and produce a "T2 weighted image".

2.2.8 Ultrafast MRI sequences: Echo-Planar Imaging

The most commonly used MRI acquisition parameters are the spin echo (SE) sequences. However, SE sequences may take up to minutes to acquire each slice. Faster sequences are therefore required for imaging dynamic processes. Although gradient recalled echo (GRE) sequences brought imaging time down to seconds (Haase *et al.*, 1986), the introduction of echo-planar imaging (EPI; Mansfield, 1977) sequences meant that it was theoretically possible to obtain a whole brain image in a fraction of a second.

The major difference between EPI and other MRI sequences is the way in which the data is sampled. Once acquired, MRI data are Fourier transformed. The Fourier transform converts the data from the time domain to the frequency domain. The two orthogonal gradients applied during frequency encoding and phase encoding mean that MRI data must be considered as lying in 2-dimensional frequency space (or k space). Typical MRI sequences sample one line of 2-dimensional k space (e.g. one single slice) after each RF pulse, whereas EPI measures all lines of k space (e.g. all slices in a volume) after a single excitation. EPI therefore greatly reduces imaging time and makes it an ideal sequence for dynamic MRI techniques such as fMRI.

All experiments in this thesis were conducted at the Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, using a Siemens VISION system (Siemens, Erlangen, Germany), operating at 2 T. This system was used to

51

acquire both T1-weighted anatomical images (1 x 1 x 1.5 mm voxels) and gradient echo planar T2*-weighted MRI image volumes with blood oxygenation level dependent (BOLD) contrast. Each echoplanar image comprised 35 1.8 mm axial slices (3 x 3 mm pixels) with a 1.2 mm slice interval, giving a resulting resolution of 3 mm.

2.3 fMRI and the magnetic properties of blood

2.3.1 BOLD contrast in fMRI

The cellular component of blood contains red blood cells (erythrocytes) which contain haemoglobin, the protein responsible for oxygen transport. Oxygen binds to iron, a constituent of the haem component of haemoglobin. When haemoglobin has no oxygen bound, each haem group has a net magnetic moment because of iron's 4 unpaired electrons (Pauling and Coryell, 1936). As soon as oxygen is bound, this net moment disappears due to a redistribution of the available electrons between iron and oxygen. The magnetic state of blood will therefore reflect its level of oxygenation.

The T2* of water protons is influenced by interactions between the protons themselves and also by local B_0 inhomogeneities caused by different magnetic properties of various molecules. Molecules such as deoxyhaemoglobin, that have a local magnetic field, are called paramegnetic. The local gradient induced by paramagnetic molecules will contribute to the decay of transversal magnetisation thereby shortening the T2* decay time of the surrounding protons. Hence changes in the levels of deoxyhaemoglobin (more precisely, changes in the ratio of deoxyhaemoglobin to oxyhaemoglobin) should result in changes in T2*. This effect was demonstrated empirically *in vivo* in animal work carried out by Ogawa and colleagues (Ogawa and

Lee, 1990; Ogawa *et al.*, 1990) and Turner and colleagues (1991). Both groups showed that experimental manipulation of the oxy- to deoxyhaemoglobin ratio (usually by hypoxia) produced detectable contrast changes in blood vessels and, critically, within the tissue water surrounding vessels. Ogawa and colleagues (1992) and Kwong and colleagues (1992) went on to demonstrate that the difference in T2* signal produced by deoxyhaemoglobin concentrations *in vivo* in humans was sufficient to act as a contrast source. This signal source was termed the blood oxygenation level dependent (BOLD) contrast.

2.3.2 Neurophysiology and BOLD

Since the early demonstrations of BOLD fMRI, the relationship between increases in neuronal electrical activity and change in blood oxygenation has not been fully understood (see Villringer and Dirnagl, 1995 for review). In Chapter 4, this relationship is explored with a number of biophysical simulations which used a haemodynamic input-state-output model. At present, the following causal chain of events is thought to be involved in the coupling between neuronal activity and BOLD signal. Task-related neuronal activity increases in specific areas of brain grey matter. The rates of oxygen and glucose usage in these areas are increased (Hyder *et al.*, 1997) causing a decrease in blood oxygenation in the capillary bed supplying the neuronally active tissue (Vanzetta and Grinvald, 1999) approximately 100 milliseconds after sensory stimulation. This evokes a release of vasodilatory compounds, producing increased blood flow to, and dilation of, these capillaries (Duelli and Kuschinsky, 1993). The oxygen supply to electrically active tissue begins to exceed demand and blood oxygenation increases in

the capillaries and venules that drain them (Villringer and Dirnagl, 1995). When neuronal activity returns to baseline, blood flow also returns to baseline, but blood volume in draining venules remains elevated for 30-60 seconds after blood flow has reequilibrated (Mandeville *et al.*, 1999). Recently it has been shown that local field potentials correlate highly with BOLD activity, more than multi-unit spike rates (Logothetis *et al.*, 2001). This suggests that the BOLD response reflects the cortical afferents and signalling within a local population area.

The neurovascular response to a transient change in metabolic activity typically involves a transient decrease in oxygenation resulting in a small, transient decrease in the BOLD response (the 'initial dip'). Next, local blood volume increases and, 500-1000 milliseconds after sensory stimulation, local blood flow increases. This causes a much larger decrease in deoxyhaemoglobin concentrations than the initial increase. The BOLD contrast increase at 5-6 seconds evoked by this deoxyhaemoglobin decrease is the signal that is typically measured in fMRI experiments. After stimulation, the deoxyhaemoglobin concentration increases and BOLD response returns to baseline at 12-15 seconds with a small undershoot that may last up to 30 seconds (Blamire *et al.*, 1992; Kwong *et al.*, 1992; Menon *et al.*, 1992). The spatial and temporal dynamics of the BOLD response may vary from person to person and from one brain region to another (Birn *et al.*, 2001; Miezin *et al.*, 2000).

The dynamics of the neurovascular response cause the BOLD signal to be delayed in time. Whereas neuronal dynamics occur in the millisecond time frame, the BOLD response takes a number of seconds to evolve. This means that although several fMRI images can theoretically be acquired in a second, the temporal smoothing of the underlying neuronal signal, effected by the BOLD response, ultimately dictates fMRI's effective temporal resolution. The time course of the neurovascular response must be accounted for in the statistical analysis of fMRI time series (see Chapter 3).

Although changes in BOLD signal are correlated with neuronal activity over a short period of time, the signal may be affected by a number of physiological and physical parameters. Physiological factors include carbon dioxide and oxygen concentrations, and the neuromodulatory environment determined in part by arousal, stress or drugs including nicotine and caffeine. Physical factors include subject and scanner temperature and magnetic field homogeneity.

As blood occupies only a small fraction of grey matter, BOLD signal changes are in the order of a few percent at best. These small signal changes require the implementation of sophisticated image processing and analysis techniques to ensure that observations reflect true BOLD signal and not noise. These techniques are described in the next chapter.

Chapter 3

Chapter 3

Methods: Statistical Analysis of fMRI Time Series

3.1 Introduction

As discussed in Chapter 1, there are two complementary principles underlying the organisation of the brain. First, different regions serve different functions, a principle known as functional specialisation. Identifying functionally specialised regions generally proceeds using statistical parametric mapping (SPM). In the studies reported in Chapters 5, 6, 7 and 8, SPM is used to investigate the neural correlates of single word reading across different ranges of design parameters. Although functional specialisation is a fundamental principle of brain organisation, it is important to recognise that there are anatomical connections amongst specialised regions that enable functional integration. Functional integration can be characterised by a number of techniques including (DCM) and Multivariate Autoregressive Modelling (MAR). In Chapter 9, SEM is used to investigate the effect of word type on functional integration amongst a number of regions involved in single word reading.

In order to assign an observed response to a particular brain structure or cortical area, the data must conform to a known anatomical space. Hence, the analysis of fMRI data starts with a series of spatial transformations of brain images aimed at reducing artifactual variance components in each voxel time series. The imaging time series is first realigned to a common reference frame to correct for subject movement during scanning. After realignment the data are transformed using linear and nonlinear mappings into a standard anatomical space (Friston *et al.*, 1995a). This normalisation procedure allows data to be averaged across subjects and permits data to be reported within a standardised reference co-ordinate system. Following this pre-processing, a statistical model must be created in order to draw inferences about changes in (a) regional brain activity and (b) functional integration between different conditions. The pre-processing and the statistical analyses of the data presented here were performed using SPM99 software (http://fil.ion.ucl.ac.uk/spm) and the SEM Toolbox of SPM99.

3.2 Pre-processing

Processing of the data prior to statistical analysis involves realignment, coregistration, normalisation and smoothing. A general scheme is shown in Figure 3.1.



Figure 3.1 Schematic representation of data pre-processing, as implemented in SPM99 software.

3.2.1 Realignment

3.2.1a Spatial

Head motion during fMRI can give rise to artifactual change in signal intensity. Despite subjects being firmly immobilised with soft head pads, even the best subjects show movement up to a millimetre or so. Realignment removes variance from a time series which would otherwise be attributable to error (hence decreased sensitivity) or to evoked effects i.e. if movement is correlated with the cognitive task. Realignment involves the following steps:

- Estimating the 6 parameters of the affine 'rigid body' transformation that minimises the [sum of squared] differences between each successive scan and the first. In three dimensions, a rigid body transformation can be defined by 6 parameters, typically three translations and three rotations about orthogonal axes.
- 2. Applying the transformation by resampling the data using sinc or trilinear interpolation.

3.2.1b Temporal

In multi-slice acquisitions, different slices will be acquired at slightly different times. In fMRI analysis, stimulus onset times are specified in scans, hence posing the problem for event-related fMRI studies that certain slices will be more sensitive to a particular model of haemodynamic responses. Temporal realignment ensures that the data from all slices within a given volume correspond to the same time point. This is achieved using sinc interpolation over time. The interpolation is unnecessary at very short TRs (e.g. <1

second). In addition, it may be inaccurate at long TRs (e.g. >6 seconds). However, the TR for the event-related fMRI study reported in Chapter 5 was 3.15 seconds and slice timing correction was therefore implemented.

3.2.2 Coregistration

Two images may be coregistered, so that they can be viewed in the same co-ordinate space. For coregistration within the same modality, SPM99 software minimises the sum of squares of the difference between target and object image.

3.2.3 Spatial normalisation

Spatial realignment produces a mean image for each subject. This mean image is used to estimate the warping parameters that map this image onto a template (in fMRI this is a template EPI image) that conforms to an international standard in terms of size and shape. The standard conforms to a normal anatomical space (in the case of SPM99 software this is the space defined with respect to the MNI-305 template, supplied by Alan Evans, MNI, Canada from the ICBM NIH P-20 project). The estimation is achieved using:

- 1. A 12-parameter affine transformation to approximate the MNI-305 template where the parameters constitute a spatial transformation matrix. The affine transformation is similar to that used during realignment but also includes zooms and shears.
- Parameters are then estimated for a set of low-spatial-frequency basis functions (a discrete cosine set) in each dimension, to optimise a smoothly nonlinear normalisation to the template.

A Bayesian framework is used to estimate the parameters, to find the deformation that is most likely given the data. The deformation is updated iteratively to minimise the sum of squared differences between the template and the deformed image and reflects the probability of actually getting that image if the transformation was correct. Prior information about the likelihood of a given transformation is incorporated by weighting the least squares (Ashburner *et al.*, 1997). This procedure can be extended to allow normalisation across image modalities, allowing, for example, the SPM obtained from a fMRI time series to be overlaid on an individual subject's structural T1 image.

3.2.4 Spatial smoothing

After normalisation, the fMRI data are smoothed by applying a 3-D Gaussian kernel (point spread function), of known width, to each voxel. The motivations for smoothing are:

- 1. It renders the data more parametric in their distribution and ensures the validity of parametric statistical tests.
- 2. It is necessary to meet the assumptions of Gaussian Field Theory (see later).
- 3. It allows data to be averaged over subjects because regional effects are expressed at a spatial scale where homologies in functional anatomy exist over subjects.
- 4. It enhances the signal to noise ratio, because the noise generally has a much higher spatial frequency than the evoked haemodynamic responses. The matched filter theorem states that the optimum smoothing kernel corresponds to the size of the effect anticipated. Optical imaging experiments suggest that the spatial scale of the haemodynamic response is about 2-5 mm.

The extent of smoothing is described by the full width half maximum (FWHM) of the Gaussian kernel. Smoothing is a linear convolution, such that the final smoothing (equivalent FWHM) is the root sum of squares of the smoothing kernels. In the experiments reported in this thesis, a smoothing of 6 mm was used, which allowed regional effects to be expressed at a spatial scale where homologies in functional anatomy exist over subjects.

3.3 Characterising Haemodynamic Responses using the General Linear Model

Following spatial pre-processing, the data are ready for statistical analysis. In the absence of prior information regarding the physical location of a particular function, statistical analysis of evoked haemodynamc responses must test for experimentally-induced effects at each intra-cerebral voxel. In the context of SPM, this involves three steps. First, the pre-processed data are fitted to a general linear model (Friston *et al.*, 1995b). Second, the parameter estimates and residual error from the general linear model are used to create statistical images (i.e. SPMs). Third, probabilistic inferences are made about the differences between experimental conditions using Gaussian Field Theory whilst controlling for the probability of false positives. A general scheme for statistical parametric mapping is shown in Figure 3.2.

Data analysis as implemented in SPM99 software is parametric. Statistics with a known null distribution are used, such that under the null hypothesis the probability of obtaining a statistic greater than, or equal to, that observed can be computed. The statistical model used is a special case of the general linear model.

61





3.3.1 Parameter estimation using the General Linear Model

Commonly used parametric models, such as linear regression, *t*-tests and analysis of variance (ANOVA) are special cases of the general linear model. This model explains variation in the data, Y, in terms of a linear combination of the explanatory variables (x), plus an error term (ϵ). For *J* observations of Y,

$$\mathbf{Y}_j = \mathbf{x}_{j1}\boldsymbol{\beta}_1 + \ldots + \mathbf{x}_{jl}\boldsymbol{\beta}_l + \ldots + \mathbf{x}_{jL}\boldsymbol{\beta}_L + \boldsymbol{\varepsilon}_j$$

The β_l are unknown parameters, corresponding to each of the L explanatory variables for the *j*th observation of Y. The errors ε_j are assumed to be independent and identically distributed normal random variables with zero mean. The general linear model can be expressed in matrix formulation:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

for the column vector of observations \mathbf{Y} , the column vector of error terms $\boldsymbol{\varepsilon}$ and the column vector of parameters $\boldsymbol{\beta}$, $\boldsymbol{\beta} = [\beta_1 \dots \beta_j \dots \beta_j]^T$. Matrix \mathbf{X} , of size $J \ge L$, is the design matrix. This matrix has one row per observation and one column per model parameter.

The number of parameters L is (usually) less than the number of observations J. This means that the simultaneous equations implied by the general linear model (obtained by expanding the matrix formulation with $\varepsilon=0$) is overdetermined (i.e. cannot be solved). Therefore, some method is required for estimating parameters that "best fit" the data. The method adopted in the context of SPM99 software is that of least squares. The values of β that minimise the sum of squares of the residual errors are known as the parameter estimates ($\hat{\beta}$). If X is of full rank, i.e. does not include linearly dependent columns, then

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{Y}$$

If X includes linearly dependant columns, it is rank deficient and there are a number of valid least-squares solutions to the general linear model. This is often the case with functional imaging studies, hence an optimal solution can only exist within certain constrains. However, if X is re-parameterised such that all explanatory variables sum to zero, the effects of a variable are estimated after the overall mean has been subtracted. When X is constrained to sum to zero, this is equivalent to using the pseudoinverse of $X^{T}X$. SPM99 software uses the Matlab implementation of the pseudoinverse. This allows calculation of a unique least squares solution, even for rank deficient design matrices:

$$\hat{\boldsymbol{\beta}} = \operatorname{pinv}(\mathbf{X}^{\mathrm{T}}\mathbf{X})\mathbf{X}^{\mathrm{T}}\mathbf{Y} = \operatorname{pinv}(\mathbf{X})\mathbf{Y}$$

3.3.2 Residual (error) variance

The residual variance σ^2 is estimated by the residual mean square, calculated as the residual sum of squares $\epsilon^T \epsilon$, divided by the degrees of freedom J-L. If the errors are

independent, equally and normally distributed across explanatory variables, then the estimated residual variance, $Est\sigma^2$, is given by

$$\text{Est}\sigma^2 = \mathbf{e}^{\mathrm{T}}\mathbf{e}/\text{J}-\text{L}\sim\sigma^2$$

In the experiments reported in this thesis, the residual (error) variance estimate is based on all scans and is voxel specific. An alternative used elsewhere involves estimation of the residual (error) variance across all voxels for any given scan/time point (Worsley and Friston, 1995).

3.3.3 T statistics

Inferences about the parameter estimates are made using their estimated variances. This allows for two types of statistical test. One can test the null hypothesis that all the estimates are zero using the *F* statistic to give a SPM{*F*} or, alternatively, that some particular linear combination or a "contrast" (e.g. a subtraction) of the estimates is zero using a SPM{*T*}. In the studies reported in this thesis, inferences about effects of interest are made using *T* statistics. An effect of interest is specified by a vector of contrast weights **c** that gives a weighted sum or compound of parameter estimates $\mathbf{c}^T \hat{\boldsymbol{\beta}}$, referred to as a contrast. An example of contrast weights could be $[1 -1 \ 0 \ 0 \ ...]$ to compare the differential responses evoked by two conditions that have been modelled by the first two condition-specific regressors (columns) in the design matrix. The *T* statistic is simply this contrast $\mathbf{c}^T \hat{\boldsymbol{\beta}}$ divided by its estimated standard error, e.g. the square root of the contrast variance:

$$T = \mathbf{c}^{\mathrm{T}} \hat{\boldsymbol{\beta}} / (\operatorname{var}(\mathbf{c}^{\mathrm{T}} \hat{\boldsymbol{\beta}}))^{1/2}$$

where the contrast variance $var(\mathbf{c}^T \hat{\boldsymbol{\beta}})$ is a function of the residual variance σ^2 and the design matrix X:

$$\operatorname{var}(\mathbf{c}^{\mathrm{T}}\,\hat{\boldsymbol{\beta}}\,) = \sigma^{2}\mathbf{c}^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{c}$$

As stated above, an important assumption in the analysis of time-series is that the residuals are identically and normally distributed. However, the haemodynamic response is of longer duration than the typical scan acquisition time, which leads to serial correlations among the error terms. In the experiments reported in this thesis, these temporal autocorrelations were estimated using a restricted maximum likelihood (ReML) and a AR(1) + white noise model (Friston *et al.*, 2002) and used to make the appropriate non-sphericity adjustment at the point of inference.

3.3.4 Global scaling

Scaling of global activity across sessions/subjects is enforced in SPM99 software. In addition, prior to parameter estimation, the global activity within a session can be scaled to a default value in order to remove scanner drifts over time. As a result, voxel specific changes in BOLD are expressed as a percentage of this global mean signal (i.e. 100).

3.3.5 High Pass Filter

Prior to parameter estimation, a high pass filter can be applied to the data in order to remove very low frequency effects such as the slow drift in BOLD signal over minutes. The filter cut off may be calculated on a study specific basis, in order to remove low frequency noise but not variation in signal due to the experimental periodicity. However, there is evidence that in most cases a high pass filter of approximately 120 seconds is sufficient to remove machine and physiological noise while preserving experimentally induced variation in signal (Zarahn *et al.*, 1997).

3.3.6 Statistical inference and the Theory of Gaussian Fields

As already described, SPM99 software calculates the significance of brain activity in a voxel-wise manner. However, inference can be drawn at a number of different levels. SPM derives p values pertaining to

- 1. *set-level inference*: the number of activated regions (i.e. number of clusters above a height and volume threshold)
- 2. *cluster-level inference*: the number of activated voxels comprising a particular region (i.e. volume of cluster)
- 3. *voxel-level inference*: the *p* value for each voxel within that cluster. Significance testing in this thesis is limited to testing significance at the voxel level.

In classical statistics, multiple tests require a correction (such as Bonferroni correction) for the number of tests performed on the data. The p values of the ensuing maxima must be corrected for multiple comparisons if the hypothesis is not constrained by anatomy (i.e. the null hypothesis states there is no effect anywhere in the brain). This correction is based on Gaussian field theory, which takes into account the fact that neighbouring voxels are not independent by virtue of spatial smoothing in the original EPI images. This correction is similar to a Bonferroni correction for multiple comparisons but generally less severe provided that the data are sufficiently smooth.

Gaussian field theory is used to generate p values corrected for multiple *dependent* comparisons in an SPM of known smoothness. This is done by considering the voxels to be a lattice representation of a continuously varying field across the brain. The application of Gaussian field theory proceeds in three steps. First, the number of resels in an image is determined. Second, based on the number of resels and the smoothing applied, the *expected* Euler characteristic of that image is computed for different statistical thresholds. Third, the *expected* Euler characteristic is used to derive the p threshold for the required control of false positives.

A resel (resolution element) can be defined as a cube of voxels of the same size as the FWHM used to smooth the data. For instance, if a smoothing of $6 \times 6 \times 6$ is applied, a resel will be a $6 \times 6 \times 6$ voxel cube (approximately). The number of resels can be computed by dividing the volume of the search region by the product of the FWHM of the smoothing kernel in each dimension (in the example above, $6 \times 6 \times 6 = 216$). In other words, the number of resels of an image depends on the number of voxels in a search region and the FWHM used to smooth the data. The number of resels in an image can be used to estimate the *expected* value of the Euler characteristic at any given threshold *T*.

The Euler characteristic of an image is given by the number of peaks in that image after it has been thresholded. Under the null hypothesis, i.e. when the distribution of T statistics is due to random noise alone, there will still be peaks and troughs in the continuously varying field across the brain. However, very high values of the statistics will be unlikely. The *expected* Euler characteristic of a threshold T is the number of peaks which are expected to be observed at that threshold, under the null hypothesis. However, at high statistical thresholds, the *expected* Euler characteristic is also a good approximation of the probability of observing one or more peaks at that threshold, under the null hypothesis. For instance, if x is the T score threshold that gives an expected Euler characteristic of 0.05, then, if the image is thresholded at x, any peaks that remain have a probability p of less than or equal to 0.05 that they have occurred by chance.

As stated above, the *expected* Euler characteristic depends on the number of resels, which is a function of the smoothness of the data. As the smoothness increases, the number of resels decreases yielding fewer independent tests, hence the probability of obtaining a peak as large as that observed gets smaller. It should be noted that the *expected* Euler characteristic depends not only on the number of resels, but also on the shape of the volume in which the resels are contained. Finally - the number of resels in an image is similar, but not identical, to the number of independent observations (if it was the same, then, instead of using Gaussian field theory for the expected Euler characteristic, we could use a Bonferroni correction for the number of resels).

The experiments reported in this thesis investigate the neural correlates of reading, which comprise a number of regions widely distributed across the brain. Voxellevel inferences are made using a statistical threshold of p<0.05 corrected for the entire volume using Gaussian field theory. However, trends at p<0.001 (uncorrected for multiple comparisons) may also be reported.

3.3.7 Inferences about subjects and populations: Random vs. Fixed effect analyses

The statistical inference drawn from fMRI time series may be of two types. First, the results may be specific to the particular subjects at the time of scanning. This fixed effect inference is drawn from the effect size averaged across subjects relative to the within subject variability. Here the degrees of freedom pertain to the number of scans across all subjects. Generally, 6 subjects are included in these analyses to provide a representative mean across subjects. The limitation of this type of analysis is that an effect size may be primarily driven by a few subjects. To overcome this one can perform a conjunction analysis across the six subjects which tests for regions commonly activated by a particular condition in all subjects. Fixed effect analysis can be thought of as an extension of a case report, commonly used in clinical studies and animal lesion experiments, where an effect is observed in a particular subject and then this effect is replicated in further subjects. The studies reported in Chapters 5, 6 and 7 used a fixed effect approach.

Alternatively, random effect analyses allow inferences to be made about the population from which the sample of subjects was drawn. One observation per subject per condition is entered into a random effect analysis (usually a contrast of parameter

70

estimates $\mathbf{c}^{\mathrm{T}} \hat{\boldsymbol{\beta}}$ from the fixed effect or 1st level analysis). Hence the effect size is compared to the between subject variability in these contrasts. This type of analysis is, therefore, not at risk of being biased by strong effects in a subset of subjects. It follows that more subjects are required to achieve a significant result with random effect analyses, as the degrees of freedom depend on the number of subjects scanned, a suitable minimum number of subjects being 12. Random effect analyses are typically one-sample *t*-tests which assess whether the estimated effect size (i.e. contrast) is significantly greater than zero across all subjects. The study reported in Chapter 8 used a random effect approach.

3.4 Characterising Effective Connectivity using Structural Equation Modelling

3.4.1 Functional and Effective Connectivity

Complex cognitive functions, such as reading processes, are likely to be mediated by a network of inter-dependent regions rather than by isolated modules. Functional integration refers to the interaction among functionally specialised brain regions and can be characterised in terms of functional or effective connectivity (e.g. McIntosh and Gonzalez-Lima, 1994; Horwitz and Sporns, 1994; Friston *et al.* 1993a; 1993b). Functional connectivity is defined as the temporal correlations among neurophysiological events in different neural systems, whereas effective connectivity is defined as the influence that one neural system exerts over another. In Chapter 7 of this thesis, effective connectivity during single word reading is investigated using structural equation modelling.
3.4.2 Structural Equation Modelling

Structural Equation Modelling is a method to analyse system-level interactions (Büchel and Friston, 2000; Horwitz *et al.*, 1999; Büchel and Friston, 1997; McIntosh and Gonzales-Lima, 1994). Before neuroimaging techniques became available, SEM has been used in social and economic sciences for several decades.

Structural equation modelling does not itself produce a model of regional interactions in the brain. Rather, SEM of functional imaging data identifies connection strengths that best predict the variance-covariance structure of the empirical data (*i.e.* fMRI time series), under the constraints of a pre-specified anatomical model. The anatomical model is specified by the researcher on the basis of neuroanatomical and neuroimaging literature. Clearly, the specified anatomical model is unlikely to be a complete account of all regions involved in the task and all possible connections amongst them. However, simplified models may be sufficient to address key questions regarding the functional or effective connectivity within a system. For each region of the anatomical model, the fMRI time series is extracted for each subject. Since functional homologous regions may vary slightly in location from subject to subject, the specific co-ordinates for these regions are selected for each subject independently. Regional activity can be defined as the first eigenvariate of all voxels in any region.

3.4.3 Model estimation

The coupling amongst regions within the anatomical model is estimated based upon correlations among regional activities over time. Generally, this is achieved by using iterative methods such as maximum likelihood or generalised least squares. Iteration

72

begins by assigning a set of start values to the free parameters of the anatomical model, from which an implied variance-covariance structure is computed and compared to the variance-covariance structure of the empirical data. Iteration continues until the difference between the implied and the observed variance-covariance structure cannot be minimised any further. This produces a single number (i.e. value of the fitting function) that is a function of the discrepancy between the two variance-covariance structures. In the context of multivariate normally distributed variables, such discrepancy is asymptotically distributed as chi square (χ^2) with (p/2)(p+1)-q degrees of freedom (where p is the number of observed variables and q is the number of free parameters (Bollen, 1989). Chi square values are derived by multiplying the value of the fitting function by the sample size minus one. Chi square values are therefore an index of how "badly" the pre-specified anatomical model fits the data and can be used to compare alternative models.

Statistical inferences about the path coefficients are based on the comparison of a free model with a constrained model. In a free model, all parameters are optimised using the minimum of the maximum likelihood function. In a constrained model, the path coefficients for one or more connections are set to have a particular value (e.g. zero). The difference in goodness of fit of the free and constrained models is given by the difference in the chi-squared values, with n degrees of freedom where n is the difference in free parameters. Under the null hypothesis, that one area has no influence over another, the free and constrained models do not differ in goodness of fit. Generally, inferences are made at p<0.05.

3.4.4 Moderator variables

To assess task-related changes in coupling, a moderator variable can be included which models the influence of the task on the coupling between two areas (Büchel and Friston, 1997). A moderator variable, or interaction term, can be constructed by multiplying the mean-centred time series of the source region by the experimental variable of interest (e.g. a vector that encodes the presentation of words *versus* pseudowords). Activity in any "target" region is then characterised in terms of a linear combination of (i) activity in the source region; (ii) the design variable; and (iii) the interaction term. The influence of the interaction term on the "target" region corresponds to the effect of word type on the coupling between the "source" and "target" regions, as depicted in Figure 3.3. The moderator variables or interaction terms play exactly the same role as psychophysiological interactions (PPIs) in simple regression analyses of inter-regional coupling (Friston *et al.*, 1997).



Figure 3.3 The model for the modulatory effect of word type on connectivity between a source and a target area (left) and its implementation in the context of SEM (right). The influence of the interaction term on the target region corresponds to the effect of word type on the coupling between the "source" and target regions.

3.4.5 Interpretation of path coefficients

There are two interpretations of significant positive values in a structural equation model. First, the activity of the source area influences the activity of the target area more under one experimental condition than another. An alternative explanation is that activity in the source area changes the extent to which the experimental condition influences activity in the target area. Both interpretations are valid and are identical to psychophysiological interactions as defined for neuroimaging (Friston *et al.*, 1997).

The dimension and units of a path coefficient depend on whether it refers to a functional connection or the moderating effect of task. For functional connections, a path coefficient of +1 indicates that a unit increase (in standard deviations of BOLD signal change) in the source area causes a unit increase in the target area. A path coefficient of 0 indicates no overall coupling between two regions. A path coefficient of -1 indicates that a unit increase (in standard deviations of BOLD signal change) in the source area causes a unit deviations of BOLD signal change) in the source area causes a unit increase (in standard deviations of BOLD signal change) in the source area causes a unit decrease in the target area. The path coefficient of a moderator variable indicates the difference in this inter-regional covariance under the different experimental conditions. In other words, the path coefficient of a moderator variable describes the extent to which the coupling between a source and a target region is influenced by the moderator variable itself.

3.4.6 Implementation of SEM

A simple structural equation model comprises p regions $x_1, x_2 \dots x_p$, from which N observations are recorded. If x is the N x p matrix of deviation (from the mean)

scores, then the observed variance-covariance structure S can be characterised as follows:

$$\mathbf{S} = (1/(N-1)) \mathbf{x}^{\mathrm{T}} \mathbf{x}$$

S is a symmetric matrix, with sample variances along the diagonal and inter-regional covariances off the diagonal. The variables $x_1, x_2 \dots x_p$ may be influenced by a set of independent variables $z_1, z_2 \dots z_p$ outside the model. This is equivalent to a model with residual influences z ($z_1, z_2 \dots z_p$). In addition, unidirectional influences between the elements of x can be specified. Algebraically, the model for x is:

$$\mathbf{xI} = \mathbf{xB} + \mathbf{z}$$

where **B** is a matrix of unidirectional path coefficients and **I** is a $N \times N$ identity matrix. Here x appears on both sides of the equation. This reduces to:

$$\mathbf{x} = \mathbf{z}(\mathbf{I} - \mathbf{B})^{-1}$$

When looking at the variance-covariance structure implied by the model and omitting the denominator 1/(N-1), we have:

$$\mathbf{x}^{\mathrm{T}}\mathbf{x} = (\mathbf{z}(\mathbf{I} \cdot \mathbf{B})^{-1})^{\mathrm{T}}(\mathbf{z}(\mathbf{I} \cdot \mathbf{B})^{-1})$$
$$= ((\mathbf{I} \cdot \mathbf{B})^{-1})^{\mathrm{T}}\mathbf{C}(\mathbf{I} \cdot \mathbf{B})^{-1}$$
$$= \mathbf{\Sigma}$$

where $\mathbf{C} = \mathbf{z}^{T}\mathbf{z}$ is the diagonal matrix of residual variances, and Σ is the variancecovariance matrix implied by the model when the denominator is omitted. Parameters in **B** and **C** are free parameters estimated by the iterative minimisation of the differences between the observed **S** and the estimated Σ variance-covariance structures. Random starting estimates are generated, from which gradient descent methods estimate the optimal parameters according to the objective maximum likelihood function (Higham, 1993; Bollen, 1989).

In this chapter, I have described the spatial and temporal transformations which need to be applied to fMRI images *prior to* statistical analysis and two complementary analytical procedures which can be used to investigate changes in regional haemodynamic responses and functional integration. These techniques were used to characterise experimentally-induced effects in the studies reported in the next chapters.

Chapter 4

Nonlinear Coupling Between Evoked rCBF and BOLD Signals: A Simulation Study of Haemodynamic Responses

4.1 Introduction

In this chapter, I consider the effects of experimental parameters on nonlinearities in the BOLD signal. A series of biophysical simulation studies is presented that explore the coupling of the haemodynamic response with the underlying neural activity and how this coupling depends on various parameters pertaining to experimental design. The biophysical simulation studies are based on an input-state-output haemodynamic model with input being stimulus/synaptic activity and output being the haemodynamic responses.

4.2 The Haemodynamic Model

The model is derived by combining the Balloon/Windkessel model (Buxton *et al.*, 1998), which specifies the coupling of perfusion to BOLD signal, with a dynamical model of the transduction of neural activity into perfusion changes (Friston *et al.*, 2000).

4.2.1 The balloon component

The Balloon component links regional cerebral blood flow (rCBF) to the BOLD signal as described in Buxton *et al.* (1998). All variables are expressed in normalised form, relative to resting values. The BOLD signal is taken to be a static nonlinear function of

Chapter 4

normalised total deoxyhaemoglobin voxel content (q), normalised venous volume (ν), resting net oxygen extraction fraction by the capillary bed (E_0) and resting blood volume fraction (V_0)

$$y(t) = V_0 (k_1(1-q) + k_2(1-q/\nu) + k_3(1-\nu))$$

$$k_1 = 7E_0$$

$$k_2 = 2$$

$$k_3 = 2E_0 - 0.2$$
(1)

This signal comprises a volume-weighted sum of extra- and intra-vascular signals that are functions of volume and deoxyhaemoglobin content. The latter are the state variables. The rate of change of volume is simply

$$\tau_0 \dot{v} = f_{in} - f_{out}(v) \tag{2}$$

Eq(2) says that volume changes reflect the difference between inflow f_{in} and outflow f_{out} from the venous compartment with a time constant τ_0 . This constant represents the mean transit time (*i.e.* the average time it takes to traverse the venous compartment or for that compartment to be replenished) and is V_0/F_0 where F_0 is resting flow. Note that outflow is a function of volume. This function models the balloon-like capacity of the venous compartment to expel blood at a greater rate when distended and includes a single parameter α based on the Windkessel model

$$f_{out}(v) = v^{1/\alpha} \tag{3}$$

where $1/\alpha = \lambda + \beta$. (c.f. Eq(6) in Mandeville *et al.*, 1999). $\lambda = 2$ represents laminar flow whereas $\beta > 1$ models diminished volume reserve at high pressures and can be thought of as the ratio of the balloon's capacitance to its compliance.

The change in deoxyhaemoglobin \dot{q} reflects the delivery of deoxyhaemoglobin into the venous compartment minus that expelled (outflow times concentration)

$$\tau_{0}\dot{q} = f_{in} \frac{E(f_{in}, E_{0})}{E_{0}} - f_{out}(v)q / v$$
(4)

where $E(f_{in}, E_0)$ is the fraction of oxygen extracted from the inflowing blood. This is assumed to depend on oxygen delivery and is consequently flow-dependent. A reasonable approximation for a wide range of transport conditions is (Buxton *et al.*, 1998)

$$E(f_{in}, E_0) = 1 - (1 - E_0)^{1/f_{in}}$$
⁽⁵⁾

The second term in Eq(4) represents an important nonlinearity: the effect of flow on signal is largely determined by the inflation of the balloon, which results in an increase of $f_{out}(v)$ and clearance of deoxyhaemoglobin increasing the BOLD signal. This effect depends upon the concentration of deoxyhaemoglobin such that the clearance attained by the outflow will be severely attenuated when the concentration is low (*e.g.* during the peak response to a prior stimulus).

In summary, the Balloon component assumes that substantial nonlinearities enter when rCBF changes are translated into BOLD changes (Buxton *et al.*, 1998). This component has only three unknown parameters that determine the dynamics: resting oxygen extraction fraction (E_0), mean transit time(τ_0) and a stiffness exponent (α) specifying the flow-volume relationship of the venous balloon. The only thing required, to specify the BOLD response, is inflow.

4.2.2 The rCBF component

The rCBF component relies on the generally accepted assumption that, over normal ranges, blood flow and synaptic activity are linearly related. Under the constraint that the dynamical system linking synaptic activity and rCBF is linear, the model is

$$\dot{f}_{in} = s \tag{6}$$

where s is some flow inducing signal. The signal is assumed to subsume many neurogenic and diffusive signal sub-components and is generated by neuronal activity u(t)

$$\dot{s} = \varepsilon u(t) - s/\tau_s - (f_{in} - 1)/\tau_f \tag{7}$$

 ε, τ_s and τ_f are the three unknown parameters that determine the dynamics of this component of the haemodynamic model. They represent the efficacy with which stimulus-related neuronal activity causes an increase in signal (ε), the time-constant for

signal decay or elimination (τ_s) and the time-constant for autoregulatory feedback from blood flow (τ_t).

The rCBF component can not account for the haemodynamic refractoriness and other nonlinearities observed in BOLD responses (i.e. has only zeroth and first order kernels). This means that, although nonlinearities between stimulus input and rCBF are apparent in certain contexts (Miller et a., 2001; Price *et al.*, 1992), they are not addressed in the simulation studies presented below.

The haemodynamic input-state-output model is sufficient to account for the more important nonlinearities observed in evoked fMRI responses (i.e. reproduce empirically-derived Volterra kernels) (Friston *et al.*, 2000). Moreover, the parameters of the model that best account for the more important nonlinearities observed in evoked fMRI responses are all biologically plausible. These parameters, as described above, are the rCBF, signal decay, flow autoregulatory time-constants, transit time, venous compliance/capacitance and resting oxygen extraction fraction.

4.3 The Biophysical Simulations

I used the haemodynamic input-state-output model to explore the dependence of PET and fMRI response estimates on various parameters pertaining to experimental design. By comparing rCBF and BOLD, it was possible to explore the nonlinearities in BOLD signal that depend on changes in the experimental design parameters and the shape of neuronal response. Crucially, the present investigation was concerned with the estimate of the average rCBF and BOLD responses *per stimulus* or *event*, and not with the statistical efficiency with which these responses are detected.

The sensitivity of BOLD responses to stimulus input/neuronal changes was addressed by looking at the difference between rCBF and BOLD responses as functions of (a) the presence of a prior stimulus (haemodynamic refractoriness), (b) epoch length for a fixed stimulus onset asynchrony (SOA), (c) SOA for a fixed number of events, (d) stimulus amplitude, and (e) a late or "endogenous" neuronal component. Assessing the effects of these parameters on rCBF and BOLD response estimates enables one to identify the optimum parameters for maximising BOLD signal changes in neuroimaging studies and speaks to differential sensitivities when using different experimental designs, different classes of neural responses and different modalities (i.e. PET vs. fMRI).

4.3.1 (i) Haemodynamic Refractoriness

The effect of a preceding stimulus on the BOLD response to a second stimulus has already been characterised using Volterra kernels (Friston *et al.*, 1988). The response to the second stimulus was isolated by subtracting the response to the first stimulus from the response to both. The resulting BOLD response to the second stimulus was markedly attenuated relative to the response to the same stimulus presented in isolation. Although profound refractory or saturation effects were manifest at SOAs of less than 1 second, they could also be detected for SOAs of up to 4-6 seconds. The fact that the response to the second stimulus was compromised by the first suggested some refractoriness in the BOLD response.

In this simulation study, the rCBF and BOLD responses to a second stimulus in the context of the first were modelled using the haemodynamic input-state-output model. Since rCBF in this model conforms to a linear convolution of underlying stimulus input/neural activity, the rCBF response to a second stimulus will not be affected by a first stimulus. In contrast, the Balloon/Windkessel component is fully nonlinear and can, in principle, account for the refractoriness of BOLD response.

4.3.2 (ii) Epoch Length

The effect of epoch length on the estimated BOLD response can be thought of in terms of the nonlinear refractoriness of BOLD responses to a second stimulus, in the context of the first, compounded by a whole train of prior stimuli. The epoch length is a variable of great interest in box-car paradigms. In PET, the haemodynamic responses are summed over a 60-90 second scan; this time can not be manipulated because of intrinsic limitations of the measurement technique. In fMRI, on the other hand, data are acquired by sampling the haemodynamic response to each stimulus and there are no constraints on epoch length. Thus, it is important to identify the optimum epoch length for maximising signal sensitivity in fMRI studies. Previous simulations suggest that, for block designs, the BOLD signal improves dramatically above 8 seconds until it reaches a plateau at around 32 seconds (e.g. Hutton *et al.*, 1998). These simulations rely on a linear model to address design efficiency. In contrast, the model employed here is nonlinear. Using the haemodynamic input-state-output model, the rCBF and BOLD responses were modelled as functions of epoch length for a fixed SOA of 1 second. Such an SOA was chosen because this is the shortest SOA that is guaranteed to be

relatively immune from profound haemodynamic refractoriness (see previous section). While rCBF per stimulus unit will not be affected by epoch length, a relative decrease in the estimated average BOLD response with increasing epoch length was expected although the nature and degree of this nonlinear effect could not be anticipated.

4.3.3 (iii) Stimulus Presentation Rate and SOA

In the last few years there has been a growing interest in the choice of SOAs in eventrelated fMRI. For any design, shorter SOAs allow a greater number of trials or events for averaging over the same period. Short SOAs also increase the amplitude of the signal thereby increasing the likelihood that it is detected. For these reasons, very short SOAs have been proposed (Burock et al., 1998; Clark et al., 1998; Dale and Buckner, 1997). However, the study by Friston et al. (1998) indicated that the nonlinear refractoriness of BOLD responses is expressed at very short interstimulus intervals. In particular, nonlinear effects can lead to a decrease in the integrated response for SOAs of one second or less. Other studies have also shown that nonlinear effects predominate at very short SOAs (Vazquez and Noll, 1998). These findings suggest that a lower limit on SOAs will be dictated by nonlinear interactions among events. Although profound refractory or saturation effects are manifest at SOAs of less than 1 second, they may also be detected for longer SOAs. Using the haemodynamic input-state-output model, the rCBF and BOLD responses for increasing SOAs were estimated for a fixed number of events. It was expected that the average BOLD response to each event would reduce with decreasing SOAs, but the nature and degree of this nonlinearity could not be anticipated.

4.3.4 (iv) Stimulus amplitude

In the above simulations, the simulated input amplitude was kept constant. However, previous studies suggest that this variable affects both rCBF and BOLD responses. For example, Dettmers *et al.* (1996) reported a logarithmic relationship between key press force and rCBF or BOLD signals in the motor cortex. In addition, Vazquez and Noll (1998) detected a nonlinear relationship between BOLD signal and stimulus amplitude in the primary visual cortex. In the present simulation, the impact of stimulus amplitude on the estimated average rCBF and BOLD responses was estimated. In addition, BOLD sensitivity to neuronal *activations* relative to an absolute baseline (i.e. positive stimulus amplitudes) and to neuronal *deactivations* relative to the same baseline (i.e. negative stimulus amplitudes) was compared.

4.3.5 (v) The form of the evoked neural response

The input parameters discussed above relate to experimental design and assume the same form of neuronal response is evoked by each trial or stimulus. In this final simulation, the notion that neuronal responses themselves may vary, in terms of the relative contributions of early and late neural components, is addressed. This may cause differences in the estimated haemodynamic responses when BOLD or rCBF is measured. The differential effect of late endogenous components on rCBF and BOLD responses to complex stimuli was therefore investigated. Here complex stimuli refer to stimuli that elicit multiple components in electrophysiological studies. The term "endogenous component" is taken from the event-related brain potential literature, which distinguishes between two classes of evoked waveform components that follow

events (Rugg and Coles, 1995). On the one hand, there is a set of components whose characteristics (amplitude, latency, and distribution) depend on the physical properties of sensory stimuli, such as their modality and intensity. These are "exogenous components". On the other hand, there is a second set of components whose characteristics depend on the nature of the subject's interaction with the stimulus (i.e. processing or top-down effects). These are "endogenous components". The eventrelated brain potentials that occur within the first 100 milliseconds of stimulus presentation tend to be exogenous, whereas the event-related brain potentials that occur later tend to be endogenous. Irrespective of the subject's state, some stimuli tend to elicit late endogenous components more than others. For instance, in the temporal lobe, visual and auditory words are known to elicit both an exogenous component and a late endogenous component (N400) (Kutas and Hillyard, 1980). These late components are due to the complex nature of verbal stimuli that involve functional integration among a number of brain regions over time. It has been argued that the N400 is only elicited by linguistic stimuli and reflects high-level lexical processing (Van Petten and Kutas, 1990).

A critical question is how the early and late components interact to differentially affect signal based on blood flow and deoxyhaemoglobin concentration. In other words, what is the differential sensitivity of rCBF and BOLD to the endogenous component of neural responses in the presence of an exogenous component? The haemodynamic input-state-output model was used to simulate the rCBF and BOLD correlates of different neuronal responses to the same stimuli. While the exogenous component remained the same, the endogenous component was manipulated so that its size, relative

87

to the exogenous component, varied from very small to very large. It was expected that the exogenous response would act like a "prior" stimulus and compromise the ability of the endogenous component to elaborate a BOLD response. This might partially account for reported failures in detecting BOLD signal relative to PET in regions involved in language processing, namely in temporal regions where reduced signal can not be accounted for by susceptibility artifacts (Veltman *et al.*, 2000).

4.4 Methods

Estimating the rCBF and BOLD signal *per stimulus* or *event* involved two stages. First, the haemodynamic model was used to simulate rCBF and BOLD responses for single events and trains of stimuli with different epoch lengths, SOAs, amplitudes and neural responses. Second, the average response per stimulus or event was estimated.

4.4.1 Simulating rCBF and BOLD responses

The input-state-output haemodynamic model previously described by Friston *et al.* (2000) was used to simulate the rCBF and BOLD responses to changes in stimulus input/neuronal activity. The six parameters of the model and the ranges of values explored are reported in Table 4.1.

The model's parameters	Values as estimated in Friston <i>et al.</i> (2000)	Parameter ranges explored
<u>rCBF component</u> Neuronal efficacy Signal decay Autoregulation	0.54 0.86 0.41	0.05 - 1.00 0.40 - 1.00 0.25 - 1.00
Balloon component		
Transit time Stiffness parameter Resting oxygen extraction	0.98 0.33 0.34	0.10 - 1.00 0.05 - 0.70 0.01 - 0.80

Table 4.1 The six parameters of the haemodynamic model and the ranges of values explored. The values of the six parameters are based on an empirical fMRI study of single word presentation at varying frequencies (Rees *et al.*, 1998). These values were identified as those that best reproduced the empirical data in terms of the first and second order Volterra Kernels as described in Friston *et al.* (2000).

4.4.2 Estimating average rCBF and BOLD responses

After simulating the rCBF and BOLD responses using the input-state-output model, the average rCBF and BOLD response to each event were estimated. The average BOLD responses were derived assuming a linear convolution model with a canonical haemodynamic response function (HRF). The HRF assumed was based on the model's impulse response function (i.e. the response to a unit stimulus of infinitely short duration). This conforms to a conventional event related analysis of fMRI time series where the parameter estimates correspond to the average BOLD response per stimulus or stimulus component. Because only the balloon component of the model is nonlinear, the translation of neural activity into rCBF is effectively a linear convolution. Therefore

by virtue of the fact the model's impulse response function was used, the parameter estimates for the rCBF response will always be unity. In contrast, the BOLD responses may be over-estimated or under-estimated depending upon how the nonlinearities in the balloon component interact with the experimental and neuronal factors manipulated.

In summary, by using a standard response estimation procedure based upon a linear convolution model, the effect of various manipulations on the estimated BOLD response was characterised in terms of a deviation from unity. The estimated BOLD response is expressed in terms of a single parameter that scales the haemodynamic response function used in the linear model. The estimates are simply the amount of the stimulus function required to explain the observed response under linear assumptions. In other words, the estimated BOLD response is the scalar co-efficient that, when multiplied by the haemodynamic response function used to construct the regressor, best predicted the observed simulated response (in a least squares sense). Estimates of < 1 suggest that the BOLD response is being under-estimated, or is less than would have been seen had the stimuli been presented in a more optimal fashion. Estimates of >1 suggest a super-additive effect of stimuli in elaborating a BOLD response, or equivalently, an over-estimation of the response to a stimulus presented in isolation.

The estimates were obtained from a standard multiple linear regression analysis, as would be implemented in SPM99 software, using the stimulus or neural activity functions, convolved with the model's HRF, as regressors. The response variable in the simulation studies was the output of the nonlinear haemodynamic model using the same input or stimulus function embodying the nonlinear dynamics.

4.5 Results

4.5.1(i) Haemodynamic Refractoriness

The first simulation addressed the effect of a preceding stimulus on the BOLD response to a second stimulus, for an SOA of 1 second. In the first instance, responses to a pair of words presented together and in isolation were modelled. The response to a second word in the context of the first was characterised by subtracting the response to the first word from the response to both. The difference between the responses to the second word when preceded by the first and when presented alone, reflects the impact of the first word on the response to the second. The differential effect of a preceding stimulus on BOLD and rCBF responses is shown in Figure 4.1. The upper panels illustrate the simulated responses to the stimuli presented together (green solid line) and separately (green and red broken lines). The lower panels illustrate the simulated responses to a word presented alone (red broken line) and to a second word when preceded by the first (green solid line). As stipulated by the model, the rCBF response to a second stimulus is not affected by previous stimulation. In contrast, it can be seen that the BOLD response to the second stimulus is markedly attenuated, with an augmented undershoot, in relation to the response observed when the stimulus is presented in isolation. In other words, the response to a stimulus is modulated by preceding stimuli to give a BOLD refractoriness.



Figure 4.1 Top panels: Simulated BOLD (left) and rCBF (right) responses to a pair of words (bars) one second apart, presented together (green solid line) and separately (red and green broken lines). Lower panels: Simulated BOLD (left) and rCBF (right) responses to the second word when presented alone (red broken line as above) and when preceded by the first (green solid line). The latter obtains by subtracting the response to the first word from the response to both. The difference reflects the effect of the first word on the response to the second. It can be seen that the BOLD response to the second stimulus is markedly attenuated in relation to the response observed when the stimulus is presented in isolation.

4.5.2 (ii) Epoch Length

This simulation addressed the rCBF and BOLD responses as functions of epoch length for a fixed SOA of 1 second. Thirty-two stimulus functions were created composed of one epoch with a length varying from 1 to 32 stimuli. Again, the rCBF and BOLD responses were simulated using the haemodynamic model which estimated the magnitude of the responses under linear assumptions. Figure 4.2 shows the differential effect of epoch length on rCBF and BOLD. It can be seen that the average BOLD response per event (solid line) falls with increasing epoch length and then levels out when length is around 10 seconds. In short, when the SOA is 1 second, the average BOLD signal per event decreases even when the average rCBF response (broken line) is not affected by previous activation or epoch length.



Figure 4.2 The differential effect of epoch length on estimated rCBF and BOLD responses per stimulus. The BOLD response per event (solid line) falls dramatically with increasing epoch length and then levels out at around 10 seconds. On the other hand, the rCBF response (broken line) is not affected by the epoch length.

4.5.3 (iii) Stimulus Presentation Rate and SOA

The third simulation addressed how the nonlinear refractoriness illustrated in the previous section depends on the interstimulus interval or presentation rate, by looking at the effect of increasing the SOA on BOLD and rCBF estimates for a fixed number of events.

A number of stimulus functions were created, each containing 8 stimuli, with an SOA varying between 0.25 and 16.25 seconds. The rCBF and BOLD responses to each stimulus function were simulated. The average BOLD response to a single stimulus was estimated as described above. Figure 4.3 shows the estimates of rCBF and BOLD responses as functions of the SOA. As stipulated by the input-state-output haemodynamic model, the rCBF signal (broken line) is not modulated by the SOA. In contrast, the estimated BOLD response (solid line) increases dramatically with SOA, until it reaches a maximum around 8 seconds. As the SOA increases further, the BOLD signal decreases slightly until it reaches a plateau for an SOA of around 12 seconds. These data suggest that the estimated BOLD response to a stimulus or event dramatically decreases with increasing presentation rate and that the optimal SOA for a single event occurs at around 8 seconds. Thus, the response to a stimulus is modulated by preceding stimuli to give a nonlinear refractoriness that depends on the interstimulus interval or presentation rate.

Chapter 4



Figure 4.3 The estimated rCBF and BOLD responses per stimulus as functions of the SOA. It can be seen that the estimated BOLD response increases dramatically with SOA until it reaches a maximum for an SOA of around 8 seconds. As the SOA increases further, the estimated BOLD response decreases slightly until it reaches a plateau for an SOA of around 12 seconds. On the contrary, the estimated rCBF response (broken line) is not modulated by the SOA.

4.5.4 (iv) Stimulus amplitude

Here the evoked rCBF and BOLD responses to a stimulus with varying amplitude were simulated. Thirty-two stimulus functions were used, each comprising a single stimulus whose amplitude varied from -4.5 to 4.5. Positive stimulus amplitudes correspond to stimuli that elicit an increase in neuronal activity relative to an absolute baseline. On the other hand, negative stimulus amplitudes correspond to stimuli inducing a decrease in neuronal activity relative to baseline. In the latter case, the presentation of each stimulus transiently interrupts the ongoing baseline activity. Negative neuronal input has a simple

and explicit meaning in the context of the model because baseline activity corresponds to the situation where the input is zero. When input falls below zero this corresponds to a decrease from resting or tonic levels of neural activity. Transient neuronal deactivations are easily modelled using negative spikes of varying amplitude. In contrast to the other simulation studies, the behaviour was highly dependent on the haemodynamic model parameters, particularly resting oxygen extraction fraction (E_0) . Thus, these simulations are reported for values of E_0 varying between 0.11 and 0.55. The left panel of Figure 4.4 shows the BOLD response as a function of stimulus amplitude and E_0 . It can be seen that the estimated BOLD response increases linearly with positive stimulus amplitude, but when stimulus amplitude is negative some highly nonlinear behaviour ensues. This nonlinear behaviour is more pronounced for relatively high values of E_0 and corresponds to a rectification-like effect. The use of the term "rectification" comes from the electrophysiology of neurons. Rectifying currents result in nonlinear voltage-current relationships very reminiscent of the stimulus response characteristics shown by the model. In the context of the present simulation study, the use of the term "rectification" indicates a partial rectification. Specifically, rectification refers to the selective attenuation of negative signal components. It can also be seen that the increase in BOLD response with stimulus amplitude shows a saturation effect. The right panel of Figure 4.4 shows the simulated rCBF response as functions of stimulus amplitude and E_0 , demonstrating how these responses are linearly dependent on stimulus amplitude irrespective of the value of E_0 . These results suggest that PET and fMRI may be differentially sensitive to decreases relative to increases in neuronal activity.



BOLD and rCBF estimated responses as functions of stimulus amplitude and E_0

Figure 4.4 BOLD (left panel) and rCBF (right panel) responses as functions of stimulus amplitude and resting oxygen extraction fraction (E_0) . It can be seen that the estimated BOLD response increases linearly with positive stimulus amplitude, but when stimulus amplitude is negative some highly nonlinear behaviour ensures. This nonlinear behaviour is more pronounced for relatively high values of E_0 . On the other hand, the rCBF responses are linearly dependent on stimulus amplitude irrespective of the value of E_0 . It can also be noticed that the increase in BOLD response with stimulus amplitude shows a saturation effect relative to the increase in rCBF response.

4.5.5 (v) The form of the evoked neural response

This simulation addressed the effect of differential neuronal responses, in terms of the relative contributions of early and late neural components, on BOLD and rCBF responses. Sixteen neuronal response functions were used, each of them comprising an early exogenous and a late endogenous component. While the exogenous component remained the same, the endogenous component was manipulated in such a way that its size relative to the exogenous component varied from very small to very large (see upper panel of Figure 4.5). The timing of the endogenous component was based upon the ERP literature where pronounced cognitive positivities (e.g. P300) and negativities (e.g. N400) represent important and robust electrophysiological correlates of late neuronal components. The input-state-output haemodynamic model was used to simulate the rCBF and BOLD responses to each neuronal response function and calculate the BOLD parameter estimates as above. The relative size of the late endogenous component differentially affected the signal generated from blood flow (rCBF) and deoxyhaemoglobin concentration (BOLD), as shown in Figure 4.5 (lower panel). A decrease in the BOLD parameter estimate (solid line) was incurred by increasing the relative size of the endogenous component. Thus, the exogenous response acted like a "prior" stimulus and compromised the ability of the endogenous component to elaborate a BOLD response. On the contrary, as enforced by the haemodynamic model, the rCBF parameter estimates (broken line) were unaffected by the relative size of the endogenous component (i.e. increased in direct proportion to the size of the late component).



Figure 4.5 Upper panel: Neuronal activity functions constructed to calculate the rCBF and BOLD parameter estimates. While the exogenous component remained the same, the endogenous component was manipulated in such a way that its size relative to the exogenous component varied from very small to very large. The early and late components were both modelled with gamma functions peaking at 100 and 400*ms* respectively. Lower panel: rCBF and BOLD parameter estimates as functions of the relative size of the late endogenous component. A decrease in the BOLD parameter estimates (solid line) occurred with increasing relative size of the endogenous component. On the contrary, the rCBF parameter estimate (broken line) was not affected by the relative size of the endogenous component.

4.6 Discussion

The biophysical simulations presented above investigated the coupling among stimulus input/neural activity, haemodynamic responses, and changes in fMRI signal. This sort of investigation is important for a better understanding of how the underlying neuronal activity may elicit BOLD signal changes. In the first instance, the dependence of PET and fMRI signals on various parameters pertaining to experimental design was explored. Then the notion that the neuronal responses themselves may differ, in terms of their amplitude and the relative contributions of early and late neural components, was addressed. The simulations were based on an input-state-output haemodynamic model whose validity, in relation to empirical nonlinear characterisations of evoked responses in fMRI and other neurophysiological constraints, has been established previously (Friston *et al.*, 2000).

4.6.1(i) Haemodynamic Refractoriness

Consistent with previous findings (Friston *et al.*, 1998), the response to a second stimulus was found to be compromised by the first, evidencing a nonlinear refractoriness of the BOLD response (see Figure 4.1). This suggests that the nonlinear behaviour of the Balloon/Windkessel component can, in principle, account for the refractoriness of BOLD response.

4.6.2 (ii) Epoch Length

As a consequence of the refractoriness of BOLD response, the average BOLD signal per event decreases with epoch length (see Figure 4.2). However the results showed that

there is only little gain, in terms of BOLD responsiveness, by limiting the epoch length to 10 seconds instead of 20 or 30 seconds. In contrast, a considerable gain would be obtained by limiting the epoch length to 1 or 2 seconds. Nevertheless, it is important to keep in mind that the results are concerned with the rCBF and BOLD response estimates per stimulus and not with the statistical efficiency with which these responses are estimated. This means that the choice of a very short epoch length may be ideal in terms of minimizing refractoriness, but is not optimal in terms of design flexibility and statistical power. For example, with epochs containing more than 2 or 3 stimuli, the haemodynamic response reaches a steady state, which increases the likelihood that the signal will be detected irrespective of its size and shape and the precision of sampling. Generally, to maximise efficiency, epoch length should approximate the time constants of the HRF (see Chapter 5). Furthermore, neurophysiological effects (i.e. the influence that one experimental parameter exerts on neuronal response) need to be considered when choosing the epoch length. For instance, the neuronal activity associated with task performance may be modulated by habituation or changes in the attentional set over time.

4.6.3 (iii) Stimulus Presentation Rate and SOA

Another effect of interest was that the estimated BOLD response to each stimulus decreased dramatically with increasing presentation rate (see Figure 4.3). This illustrates that haemodynamic refractoriness depends on the stimulus onset asynchrony or rate. Previous studies suggested that the optimal fixed SOA for BOLD responsiveness for a single event was 12-16 seconds (Bandettini and Cox, 1998; Hutton *et al.*, 1998). These

simulations relied on a linear model and therefore could not account for the nonlinearities observed in evoked fMRI responses. By contrast the simulation presented here, which relies on a fully nonlinear model and accounts for the nonlinearities observed in real time series, suggests that an average SOA of 8 seconds may be better because successive stimuli interact in a super-additive fashion. This can be attributed to the fact that, during the flow undershoot following a first stimulus, deoxyhaemoglobin concentration is greater than normal, thereby facilitating clearance of deoxyhaemoglobin following a second stimulus (Friston et al., 2000). What are the implications of these findings for experimental design? These results might suggest that an SOA of 8 seconds minimizes BOLD refractoriness and thus the choice of such a long SOA is optimal. However, this is not the case because a long SOA (a) severely restricts the latitude of experimental paradigms for event-related fMRI and (b) greatly reduces the design efficiency by limiting the variability in the signal or, equivalently, by reducing the low frequency components that are "passed" by the HRF (Josephs and Henson, 1999; Paradis et al., 1998). Thus, the choice of the SOA should take into account both the maximisation of the BOLD responsiveness given by a long SOA and the statistical power afforded by a short SOA. Miezin et al. (2000) reported that, despite BOLD refractoriness effects occurring in motor and visual cortices, the statistical power for detecting a response was significantly greater at fast rates owing to the increased number of trials possible and the increased signal-to-noise ratio. However, the results of the present simulation studies suggest that at some lower limit on the SOA the decreased response to each stimulus may supervene over increases in design efficiency. Finally,

Chapter 4

neuropsysiological effects (i.e. the influence that one experimental parameter exerts on neuronal response) need to be considered (see Chapter 6).

4.6.4 *(iv) Stimulus amplitude*

Stimulus amplitude was also manipulated to investigate its effect on average rCBF and BOLD responses per event. Results suggest discrepancy between PET and fMRI responses to stimuli of varying amplitude (see Figure 4.4). In particular, in relation to the absolute "resting" state afforded by the haemodynamic model formulation, fMRI appears to be asymmetric in its ability to detect deactivations relative to activations, again due to the inherent nonlinearities in BOLD signal transduction. This nonlinearity can be thought of as haemodynamic "rectification". This may account for reported failures to detect negative effects of stimulus rate when brain activity is measured with fMRI as opposed to PET (see Chapter 6; Mechelli *et al.*, 2000). The increase in BOLD response with stimulus amplitude also showed a saturation effect, which can be explained in terms of haemodynamic refractoriness. In short, findings reported here suggest that PET and fMRI may be differentially sensitive to decreases in neuronal activity.

4.6.5 (v) The form of the evoked neural response

A critical aspect of the present work was the finding that varying neuronal responses themselves may have differential effects on rCBF and BOLD signals. Here early components refer to evoked electrophysiological changes that occur within 100 milliseconds. Late components refer to those wave form components that show clear cognitive effects such as the P300 and N400. Interestingly, intermediate components such as face specific N200 are insensitive to a wide variety of top-down processes (Puce et al., 1999). Results showed that, when the relative size of a late neuronal component increased, there was a decrease in the BOLD parameter estimates even though the rCBF estimates remained constant (see Figure 4.5). Such differential responses may account for the reported failures to detect BOLD signal in temporal regions involved in language processing that are not subject to susceptibility artifacts (see Veltman et al., 2000). Verbal stimuli involve functional integration among a number of brain regions over time and therefore elicit very large late endogenous components (Kutas and Hillyard, 1980). The findings reported here suggest that BOLD sensitivity decreases when the relative size of the late endogenous neuronal component is increased and therefore the signal reduction observed in some regions of the temporal lobe may be related to the complex nature of verbal stimuli. In this context, the reduced BOLD response to the late endogenous component can be thought of as an effect of prior stimulation given by the exogenous component. In other words, the decrease in BOLD response to increasing the late endogenous component can be explained in terms of BOLD refractoriness. Interestingly, in many other parts of the brain it may not be possible to elicit an early component without inducing a late component a few hundred milliseconds later. This suggests that there may be a regionally specific variation in the sensitivity of BOLD responses to the total amount of neuronal activity that is a function of top-down effects intrinsic to particular brain regions and functions.

4.7 General Discussion

The main effects of preceding stimulation, SOA, epoch length, stimulus amplitude, and the relative size of the late endogenous neuronal component on rCBF and fMRI signals have been reported and discussed. Simultaneous manipulations of the above parameters were also performed in order to address their interactions and find quantitative, as opposed to qualitative, variations in the effect of one parameter when changing the others. These interactions were fully expected since the effects of the above parameters share a common origin, namely BOLD refractoriness. Overall, the findings reported here (i) support the idea that varying parameters that specify the experimental design may have differential effects in PET and fMRI, (ii) suggest that fMRI is asymmetric in its ability to detect deactivations, relative to activations, and (iii) suggest that relative insensitivity to BOLD signal in specific regions, such as temporal areas, may be partly explained by complex stimuli eliciting a large late endogenous neuronal component.

Clearly the relevance of the nonlinear behaviours shown here rests upon the validity of the Balloon/Windkessel component of the haemodynamic model used. Its validity was partially established in Friston *et al.* (2000) in which the model parameters required to reproduce empirically observed nonlinear responses fell exactly into plausible biological ranges. However, the nonlinear behaviours reported here, for example haemodynamic rectification, should be regarded as proof of concept under the assumption that the model employed was sufficient to emulate real haemodynamics at the level modelled. It is quite possible that the phenomena described may be trivial in certain domains of parameter space or under certain experimental conditions.

Furthermore, the model does not account for nonlinearities that enter between stimulus input and rCBF, only those entering between rCBF and BOLD.

Although the investigation reported here provides theoretical evidence for refractoriness and rectification effects in fMRI, the generalisation of these findings to all situations should be cautioned. For example, neuronal activity and/or its coupling to BOLD signal differ across different brain regions. This is supported by the fact that the amplitude and timing of the BOLD response in one region fails to predict those properties in another region, even for within-subject comparisons (Miezin *et al.*, 2000). However, a number of studies suggest that refractory effects occur not only in the superior temporal lobe (Friston *et al.*, 1998), but also in other parts of the brain such as somatosensory, motor and visual cortices (Miezin *et al.*, 2000; Cannestra *et al.*, 1998).

In summary, this chapter presents a theoretical contribution on the coupling of the haemodynamic response with the underlying neural activity and how this coupling depends on various parameters pertaining to experimental design. Empirical characterization of the nonlinearities described by the present study is required to further validate the findings and the model itself.

Chapter 5

Estimating Efficiency a priori: a Comparison of Blocked and Randomised Designs

5.1 Introduction

This chapter is about the *a priori* estimation of efficiency of fMRI designs. Efficiency of response estimation is a measure of the reliability with which model parameters are estimated, and affects the sensitivity with which experimental effects are detected. Friston *et al.* (1999) presented a mathematical framework to compute the expected efficiency for various classes of experimental designs and compared contrasts testing for (i) evoked responses *per se* and (ii) differential evoked responses among trial types. The results showed that efficiency of response estimation for a particular design depends on the contrast tested.

Here I revisit the relative efficiencies of different contrasts in the context of blocked and randomised designs in event-related experiments. Efficiency of response estimation is inversely related to the estimator variance (see Appendix). The estimator variance factorises into the error variance and the design variance. The error variance (σ^2) is the residual variance after evoked changes, modelled by the design matrix, have been discounted. The design variance $(\mathbf{c}^T(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{c})$ is a function of the contrast (**c**) and the design matrix (**X**) and embodies the variance of the explanatory variables and the correlations amongst them. The design variance can be thought of as the estimator variance when $\sigma^2 = 1$.
In the context of response detection using a single basis function, blocked designs are more efficient than randomised designs (Friston et al., 1999), as discussed in section 1.3 of Chapter 1. A useful way to understand why blocked presentation of trials is more efficient is in terms of the frequencies in which experimentally induced variance lies. Blocked designs typically induce greater variance in low frequency components that are "passed" by the HRF (Josephs and Henson, 1999; Paradis et al., 1998). This results in the standard error being smaller for contrasts testing for activations in blocked than randomised designs. The T values will therefore be higher in the blocked designs even if the amplitude of the haemodynamic responses and the parameter estimates $(\mathbf{c}^{\mathrm{T}}\hat{\boldsymbol{\beta}})$ are identical. This leads to the somewhat counter-intuitive prediction that significant effects can be evident in blocked but not randomised designs and yet the amplitude of the haemodynamic responses may not be significantly different. In a design that includes both blocked and randomised components, this means that it should be possible to show significant effects within the blocked but not the randomised component and yet no interaction between the effects and the experimental design. This is precisely what I show in the study presented in this chapter, in which I quantify the effects empirically and compare the standard errors of blocked and randomised designs.

The study also addresses a second issue which relates to the use of the design variance and the error variance in *a priori* estimation of efficiency. As discussed in Chapter 1, the design variance (i) depends on the contrast and design matrix only, (ii) can be computed *a priori*, and (iii) is the same across the whole brain. In contrast, the error variance (i) depends on cognitive/physiological effects (e.g. the haemodynamic responses may be more variable in one context relative to another), (ii) can only be

estimated by performing a statistical analysis, and (iii) is voxel-specific. This means that the relative efficiency of response estimation can only be quantified *a priori* by assuming that the error variance is independent of changes in the experimental design. For this assumption to be met, the responses need to conform to a linear convolution model, which embodies two further assumptions: nonlinearities can be discounted and the form of the haemodynamic response function is the same for different experimental designs. When these assumptions are violated, differences in error variance may arise, thereby compounding the relative efficiency of the designs. The study presented in this chapter investigates whether the relative efficiency of blocked and randomised designs can be predicted under the assumptions adopted generally in the context of *a priori* estimation of efficiency. Differences in the error variance would indicate that both the error variance and the design need to be considered when estimating efficiency. Alternatively, if differences in error variance are negligible, the relative efficiency can be predicted from the design alone.

In summary, here I present an event-related fMRI study of single word reading which involved acquiring data using two stimulus sequences associated with different efficiencies. The study aimed to (i) show that differences in efficiency, attributable solely to experimental design, may lead to identical responses that can be detected in one presentation mode but not the other, and (ii) test whether differences in error variance were significant. If this is the case, both the error variance and the design need to be considered when estimating efficiency.

Chapter 5

5.2 Methods

5.2.1 Subjects

Informed consent was obtained from 12 right handed healthy volunteers (3 male), aged between 20 and 38 (with a mean age of 26), with English as their first language.

5.2.2 Design

The stimuli were written words presented in lower case (Courier font) one at a time on a visual monitor with a *minimal* SOA of 1.5 seconds. To establish an interstimulus baseline, null events were included with an occurrence probability of 0.25 (Josephs and Henson, 1999), producing a *mean* SOA of 2 seconds. Subjects were instructed to fixate on a cross in the middle of the screen and read words silently as soon as they appeared. An eye-tracker was used to monitor the eye-movements of the subjects, in order to ensure that they kept their eyes open and scanned the stimuli.

Two variables were manipulated, conforming to a 3 x 2 factorial design. The first was stimulus duration: words remained on the screen for 200, 600 or 1000 milliseconds (msec). This factor was manipulated within a scanning session. The second variable was presentation sequence: blocked vs. randomised. In blocked sequences, stimuli were presented in trains of 35 of the same duration, resulting in a block length of 70 seconds (see upper part of Figure 5.1). In randomised sequences, stimuli of different duration were randomly intermixed (see lower part of Figure 5.1). The sequence factor was manipulated between scanning sessions.



Figure 5.1 Examples of blocked and randomised (i.e. event-related) presentations. In both blocked and randomised sequences, words either remained on the screen for 200, 600 or 1000 msec. In the blocked sequence 35 stimuli with the same duration were presented consecutively whereas in the randomised sequence, stimuli of different durations were intermixed randomly. In both cases, null events occurred randomly to enable comparison of stimulus effects with interstimulus baseline (i.e. fixation).

5.2.3 Stimuli

Each subject was presented with 420 regularly spelled words composed of 4, 5 or 6 letters. These words, with regular grapheme-phoneme relationships, were allocated to 12 different sets with 35 words in each set, matched across set for frequency (Kucera and Francis, 1967), length and number of syllables. Over subjects, each set of words was presented an equal number of times at each duration in both the blocked and randomised sessions.

5.2.4 Scanning Technique

A 2 T Siemens VISION system was used to acquire both T_1 anatomical volume images and T_2 *-weighted echoplanar images with BOLD contrast (see Chapter 2). The repetition time was 3.15 s/volume. With a minimal SOA of 1.5 seconds, this resulted in sampling every 15 msec of peri-stimulus time.

Data were acquired during four sessions, each comprising 74 volume images. The first six volumes in each session were discarded to allow for T_1 equilibration effects. In two of the four sessions stimulus duration was randomised. In the remaining two sessions, stimulus duration was blocked and counter-balanced across sessions (i.e. ABC in session 1 and CBA in session 2), within and between subjects. Sessions with randomised and blocked sequences were also alternated in a counter-balanced order within and between subjects.

5.4.5 Data Analysis

Data were analysed using SPM99 software (see Chapter 3). In brief, all volumes from each subject were realigned using the first as reference, spatially normalized to a standard MNI-305 template and smoothed with a 6 mm full width at half maximum isotropic Gaussian kernel.

The aim of the statistical analysis was to compare the effect of stimulus duration (1000 msec vs 200 msec) during the blocked and the randomised sequence and to establish the contribution of noise and design to the detectability of these effects. Three statistical models were required. The first involved data from all four scanning sessions in order to compare the effect of stimulus duration during blocked and randomised sequences ("Combined analysis"). The second and third statistical models involved independent analyses of blocked sequences or randomised sequences only. These "independent analyses" allowed the estimation of error variance for each type of stimulus sequence in regions showing duration effects in the combined analysis.

The data were high-pass filtered using a set of discrete cosine basis functions with a cutoff period of 512 seconds. Although this cutoff period only removed low frequency drifts, it was chosen in order to preserve the experimental variance in both the blocked and the randomised design. The temporal autocorrelations in the errors were estimated using a restricted maximum likelihood (ReML) and a AR(1) + white noise model (Friston *et al.*, 2002) and used to make the appropriate non-sphericity adjustment at the point of inference.

It should be noted that all analyses were performed in a fixed effect fashion, in which effects are averaged across subjects and compared to the within subject variability. In contrast, in a random effect analysis, the effect size is compared against the variability in the subject-specific parameter estimates, which embodies both the between subject and the within subject variability. This means that, given two studies that have the same between subject variability, the one with greater efficiency at the first level will also have greater efficiency at the second level. In this sense, fixed effect efficiency partially predicts random effect efficiency.

Combined Analysis

Here it was assumed that the error variance was the same for both blocked and randomised components of a compound design looking at the same activation. Stimuli were classified into six event-types according to presentation (blocked and randomised) and duration (200, 600 or 1000 msec). Both blocked and randomised sequences were modelled in the same "event-related" manner, using a canonical HRF. The parameter estimates were computed specifically for each event-type using conventional least squares. Contrasts included (i) the effect of stimulus duration, (ii) the effect of stimulus sequence, and (iii) the interaction between duration and sequence. The statistical threshold for the effect of stimulus duration was set at 10 or more contiguous voxels surviving a height threshold of p<0.05 (corrected for multiple comparisons).

To fully characterise the data, the *T* statistic for each identified effect was decomposed into the contrast (e.g. the numerator of the *T* statistic) and standard error (e.g. the denominator of the *T* statistic) (see Appendix). The contrast depends on the value of the parameter estimates $(\hat{\beta})$ and is an index of the size of the effect of interest. The standard error is inversely related to efficiency and can be factorised into the design

114

and error variance. Since the error variance in the combined analysis is operationally the same for the blocked and randomised sequences, identifying differences in the error variance for the different sequences required separate statistical models.

Independent Analyses

Within each mode of presentation (blocked or randomised), stimuli were classified into three event-types according to the time that they remained on the screen: 200, 600 or 1000 msec. Again the haemodynamic responses for each event-type were modelled with a canonical HRF and the parameter estimates were computed independently for each event-type using conventional least squares. Error variance from blocked and randomised data was compared using the F ratio (Sheskin, 1996), in voxels that showed positive effects of duration in the combined analysis. In order to maximise sensitivity to differences in the noise variance of the two data sets, comparisons were performed in this subset of voxels with a statistical threshold of 0.05 (uncorrected).

5.3 Results

5.3.1. Combined Analysis

Main effects of Stimulus Duration

Positive effects of stimulus duration (1000>200 msec) over blocked and randomised presentation were found in the bilateral superior occipital, superior lingual and fusiform gyrus (see Table 1 for details). These effects were associated with increased activation for reading with 1000 msec word duration relative to fixation (p<0.05 corrected for

multiple comparisons). There were no interactions between stimulus duration and sequence in any of these areas, even when the statistical threshold was lowered to p<0.01 (uncorrected). Negative effects of stimulus duration (200>1000 msec) were not found.

Simple main effects of Stimulus Duration

A number of regions, including bilateral superior occipital, bilateral fusiform, right superior lingual, left inferior lingual and right superior frontal gyrus, showed positive duration effects for the blocked mode (p<0.05 corrected) with only trends (p<0.001, uncorrected) for the randomised mode. For instance, in the left superior occipital gyrus, Z-scores associated with the blocked sequence were 6.2 and 6.1 whereas Z-scores associated with the randomised sequence were 3.5 and 3.4 (see Table 1). In addition, bilateral fusiform and right superior frontal gyrus showed positive duration effects for blocked presentation (p<0.05 corrected) but not for randomised presentation even when lowering the threshold to p<0.001 (uncorrected). For instance, in the right fusiform, the Z-score associated with the blocked mode was 7.2 whereas Z-score associated with randomised sequence interactions in any of these areas, even when the statistical threshold was lowered to p<0.01 (uncorrected). This suggests that the parameter estimates did not differ. There was no instance of a randomised effect being significant in the absence of a significant blocked effect.

Effect of Stimulus Sequence

Main effects and simple main effects of stimulus sequence were not detected, even when lowering the statistical threshold to p<0.001 uncorrected for multiple comparisons.

Interaction between Sequence and Duration

A duration by sequence interaction was only observed in the frontopolar gyrus (Z-score = 5.0; x=4 y=68 z=22). This region showed a duration effect (i.e. 1000 msec > 200 msec) greater under blocked presentation.

5.3.2 Independent Analyses

Independent analyses of blocked and randomised data sets were carried out solely to establish differences in the error variance for the different presentations. Results indicated that error variance differed significantly between blocked and randomised data sets in a number of regions which showed experimental effects in the combined analyses. Specifically, error variance was larger for randomised than blocked presentation in bilateral fusiform, left inferior lingual and right superior frontal gyrus. However, it was larger for blocked than randomised presentation in bilateral superior lingual areas (see Table 5.1 for details).

			Z-scores			Betas c ^T est(<i>B</i>)		Standard Error $(\sigma^2 \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c})^{1/2}$		Design Variance c ^T (X ^T X) ⁻¹ c		Error Variance σ ²		Non-sphericity <i>p</i> value
			B & R	В	R	В	Ŕ	B	R	B	Ŕ	В	R	P
Occipital														
R. Superior Occipital	16 -94	0	7.5	7.7	4.6	2.46	2.31	0.27	0.50	0.06	0.19	1.33	1.24	n.s
L. Superior Occipital	-22 -96	16	6.2	6.6	3.5	1.78	1.75	0.27	0.50	0.06	0.19	1.25	1.35	n.s.
	-20 -100	8	6.1	6.4	3.4	1.58	1.56	0.25	0.45	0.06	0.19	1.04	1.10	n.s.
L. Superior Lingual	-12 -94	-8	7.1	5.3	5.2	1.47	2.68	0.29	0.51	0.06	0.19	1.47	1.25	<i>p</i> <0.05 (B>R)
R. Superior Lingual	12 -94	-6	6.6	5.6	4.4	1.76	2.54	0.31	0.57	0.06	0.19	1.92	1.53	<i>p</i> <0.05 (B > R)
L. Inferior Lingual	-4 -86	-14	4.3	5.6	3.1	2.33	2.40	0.42	0.76	0.06	0.19	2.64	3.05	<i>p</i> <0.05 (R > B)
R. Fusiform	30 -78	-16	5.8	7.2	2.3	2.65	1.45	0.34	0.62	0.06	0.19	1.97	2.18	<i>p</i> <0.05 (R > B)
L. Fusiform	-30 -78	-20	5.6	7.4	2.4	2.50	1.48	0.34	0.62	0.06	0.19	1.90	2.15	<i>p</i> <0.05 (R>B)
Frontal														
R. Superior Frontal	18 66	24	3.1	5.5	0.5	1.92	0.31	0.34	0.63	0.06	0.19	1.28	1.44	<i>p</i> <0.05 (R > B)

Table 5.1 Positive effect of stimulus duration. Effects are reported that reached a threshold of p<0.05 (corrected for multiple comparisons), with an extent threshold of 10 or more contiguous voxels (shown in bold). While Z-scores, betas, standard errors and design variances for blocked (B) and randomised (R) data sets were derived from the combined analysis, the error variances were identified by performing independent analyses. Error variances for blocked and randomised data sets were compared with an F ratio at an uncorrected threshold of 0.05 (n.s. = not significant). In this context, non-sphericity refers to the ratio of error variances for blocked and randomised designs.

Chapter 5

5.4 Discussion

In this chapter, I have presented an event-related fMRI study of single word reading which aimed to (i) show that differences in design variance for blocked and randomised designs may lead to identical responses that can be detected in one presentation mode but not the other; and (ii) test whether differences in error variance were significant.

5.4.1 Differential Efficiency for Blocked and Randomised Sequences

As predicted theoretically, when the error variance is operationally fixed for blocked and randomised trial presentations, contrasts testing for blocked effects are more efficient leading to higher Z-scores. The ratio of standard errors for the blocked and the randomised contrast is about 1:1.82. Note that this ratio is the same for all regions because σ^2 is fixed for the blocked and the randomised mode. This means that the T (and Z equivalent) values are correspondingly higher for the simple effects of duration under blocked presentation. Although there is anecdotal evidence that the effects (β) may vary for randomised and blocked presentation in some regions, the absence of an interaction (even at p < 0.01 uncorrected) indicates that they are not significantly different. One of the most telling results here is a Z-score of 6.2 and 6.1 in the left superior occipital gyrus under blocked presentation and a Z-score of 3.5 and 3.4 under randomised presentation. This disparity is a reflection of, and only of, the design efficiency enjoyed by blocked presentation (see Table 1). It has nothing to do with noise or, indeed, a different haemodynamic response. In short, although responses were not significantly different for blocked and randomised presentation modes, activations could be detected at corrected level in blocked conditions that could not be found in event-related conditions.

5.4.2 Differential Error Variance for Blocked and Randomised Sequences

The results also indicate that the behaviour of error variance may depend on the experimental design. This suggests that, while it is possible to use the design variance to estimate the impact of the experimental design on efficiency *a priori*, it is not possible to assume that the impact of error variance on efficiency is negligible. In other words, both the design and error variance must be taken into account when estimating the relative efficiency of blocked and randomised modes. Such a finding may result from the violation of any of the assumptions underlying *a priori* comparison of efficiency. For instance, differences in the shape of the haemodynamic response or expression of BOLD nonlinearities may have led to differential error variance for the blocked and the randomised presentation.

The impact of the experimental design on error variance was variable across the brain. While the right superior frontal gyrus and most regions in the occipital cortex showed greater error variance for randomised than blocked presentation, the bilateral superior lingual gyrus showed greater error variance for blocked than randomised presentation. Furthermore, early visual areas including the bilateral occipital gyrus did not show differential error variance for the two presentation modes. Such variability may speak to functional specialization. For instance, in areas involved in high-order functions such as the right superior frontal gyrus, cognitive effects may affect the shape and timing of the neuronal response, which may result in differential error variance estimates for blocked and randomised sequences. More specifically, one possibility is that there is a greater neuronal response in these areas for the stimulus immediately following a change in duration. Because such duration-transitions are more common in

the randomised than blocked designs, this unmodelled effect would engender error. In contrast, in early visual areas such as the superior occipital gyrus, neuronal activity may simply reflect the amount of overall stimulation while not being affected by variables such as stimulus sequence.

5.5. Conclusion

In this chapter, I have shown that the relative efficiency of different presentation modes is directly expressed as differences in standard error that can have substantial effects on the ensuing statistics. When assessing responses in different experimental contexts, it is therefore important to examine the response by context interaction and avoid anecdotal comparison of the contrast-specific statistics. Furthermore, the results indicate that the error variance may depend on the experimental design. This makes *a priori* estimation of efficiency problematic: both the design and error variance must be taken into account but the latter can only be estimated by performing a statistical analysis.

Chapter 6

Chapter 6

The effects of presentation rate during word and pseudoword reading: a comparison of PET and fMRI

6.1 Introduction

In the study presented in this chapter, the effects of stimulus rate on neuronal activity during reading were investigated using fMRI and compared to those previously established with PET. This (i) enabled the segregation of brain regions showing differential rate-dependent responses, (ii) identified the optimum experimental design parameters for maximising neurophysiological responses, and (iii) allowed the sources of discrepancy between PET and fMRI signals to be evaluated further.

A previous PET study has found positive effects of stimulus rate during silent reading in a number of regions (Price *et al.*, 1996a). These include bilateral fusiform, precentral and superior temporal gyri, left supplementary motor area, left prefrontal cortex, bilateral putamen and bilateral medial cerebellum. In contrast, a negative effect of stimulus rate was detected only in bilateral temporo-parietal junctions. However, as shown by studies of auditory word processing, the effects of stimulus rate may differ in PET and fMRI (Rees *et al.*, 1997; Dhankhar *et al.*, 1997; Price *et al.*, 1992). This is because PET and fMRI (i) measure different physiological parameters that may not share a common stimulus rate dependency; (ii) rely on the estimation of different variance components (between-epoch *vs.* within-epoch) which may result in

differential sensitivity; and (iii) differ in the attentional set during resting baselines which may account for further discrepancies (see Chapter 1).

A question of interest is whether the differences between PET and fMRI rate dependency observed in the auditory domain are also seen during visual word processing. The effect of visual word rate has not previously been addressed with fMRI. In this Chapter, a study is presented which investigated the effect of stimulus rate on neural activity measured by BOLD contrast, during silent reading of words and pseudowords.

A first aim of this study was to attempt to replicate rate-sensitive activations observed with PET using fMRI. On the basis of the findings reported by Price *et al.* (1996a), a positive effect of rate was predicted in regions associated with visual processing (bilateral posterior fusiform) and response generation (cerebellum, supplementary motor area, and bilateral precentral gyri). However, some discrepancy due to measuring neural activity with fMRI rather than PET was also expected. For instance, it was anticipated that the rate-independent effects in the medial lingual gyrus and Wernicke's area, previously detected in PET but not in fMRI for auditory word processing, might not be replicated (Rees *et al.*, 1997; Dhankhar *et al.*, 1997; Price *et al.*, 1992).

A second aim of the study was to investigate the interaction between stimulus rate and word type (i.e. words vs. pseudowords) in regions involved in reading. Pseudowords are letter strings that are not real words, do not have semantic representations but can be pronounced on the basis of sublexical spelling to sound relationships (e.g. lenner). Previous studies have observed greater activation for

123

reading pseudowords than words, particularly in the left inferior frontal and left inferior temporal gyrus (Hagoort *et al.*, 1999; Fiez and Petersen, 1998; Herbster *et al.*, 1997; Rumsey *et al.*, 1997; Price *et al.*, 1996b). However, results are highly inconsistent across studies (see Chapter 8 for a review). Differential effects of stimulus rate on words or pseudowords in these or other areas may help to explain the discrepancies in the existing neuroimaging literature.

6.2 Methods

6.2.1 Subjects

Informed consent was obtained from 6 right handed volunteers (5 males), aged between 20 and 34 (with a mean age of 24), with English as their first language.

6.2.2 Design and Task

The experiment comprised a 3 x 2 factorial design alternated with a resting condition. The variables were: (i) stimulus presentation rate (20, 40 and 60 wpm as in Price *et al.*, 1996a), and (ii) stimulus type (words and pseudowords). The task involved silent reading of words and pseudowords as soon as they appeared on the screen. The resting condition involved fixating a cross in the middle of the screen. An eye-tracker was used to ensure that the subjects kept their eyes open and attended to the stimuli.

6.2.3 Stimuli

Each subject was presented with 210 words and 210 pseudowords. Stimuli were composed of 4, 5 or 6 letters and were presented in 12 blocks, with each of the 6

experimental conditions presented 5 times in a counter-balanced order across subjects. Each block lasted 21 seconds and was followed by 16 seconds of fixation. Words were matched for frequency (Kucera and Francis, 1967), length and number of syllables between blocks and were as regular as possible in terms of grapheme-phoneme conversion. Pseudowords were created from these words by changing the onset, the internal consonants or the coda. Examples of the words and corresponding pseudowords include: toast-noast, letter-lenner, and lemon-lenos. Words and pseudowords were matched for bigram frequency between blocks. Each stimulus was presented in lower case, Courier font, for 600 milliseconds, with an SOA of 1000, 1500 or 3000 milliseconds (for 60, 40 and 20 wpm respectively).

6.2.4 fMRI Scanning Technique

A 2 T Siemens VISION system was used to acquire both T_1 anatomical volume images and T_2 *-weighted echoplanar images with BOLD contrast (see Chapter 2). A total of 366 volume images were taken continuously with a scan repetition time of 3.15 s/volume, the first six (dummy) volumes in each session being discarded to allow for T_1 equilibration effects. Stimulus presentation was arranged so that every 90 milliseconds of peristimulus time was sampled equally over the session.

6.2.5 Data Analysis

Data were analyzed with statistical parametric mapping SPM99 software (see Chapter 3). In brief, functional images from each subject were realigned using the first as reference, spatially normalized with respect to the MNI-305 template, and smoothed

with a 6 mm full width at half maximum isotropic Gaussian kernel. To ensure that the functional and structural images were spatially aligned, the anatomical MRI was coregistered to a mean image created using the realigned volumes.

Stimuli were classified into six event-types according to stimulus presentation rate (20, 40 and 60 wpm), and stimulus type (words and pseudowords). A boxcar waveform convolved with a synthetic HRF was used as the reference waveform for each rate (categorical analysis) or modulated over conditions by presentation rate and presentation rate squared (parametric analysis). Because only three stimulus rates were used, these statistical models are identical. The second is simply a reparameterization in terms of linear and quadratic effects. Differences in global flow within subjects were removed by using proportional scaling as implemented in SPM99 software. The data were high-pass filtered using a set of discrete cosine basis functions with a cutoff period of 156 seconds. The temporal autocorrelations in the errors were estimated using a AR(1) + white noise model (Friston *et al.*, 2002) and used to make the appropriate non-sphericity adjustment at the point of inference. The data were analyzed in a fixed effect fashion for consistency with the previous PET study by Price *et al.* (1996a).

Categorical analysis

In order to address activations for reading relative to fixation for each presentation rate and stimulus type, a categorical analysis was performed. This analysis allowed each of the six conditions to be independently modeled and compared to fixation. Regions that showed increased activity for reading relative to fixation irrespective of rate were identified by looking at the effect of reading relative to fixation summed over reading conditions (p<0.05 corrected) in areas that were also more active for each stimulus rate (20 wpm, 40 wpm, 60 wpm) relative to fixation (p<0.01 uncorrected). This response profile was independently assessed for both words and pseudowords, words only, and pseudowords only.

Parametric analysis

In order to characterize the nonlinear relationship between neuronal activity and stimulus rate, a nonlinear regression was also employed (Büchel *et al.*, 1996). Here the effects of stimulus rate were modeled as a polynomial expansion of stimulus rate up to second order. This identified (i) positive linear dependencies; (ii) negative linear dependencies; (iii) positive quadratic (U-shaped) dependencies; and (iv) negative quadratic (inverted U-shaped) dependencies. Each of these response profiles was assessed for both words and pseudowords, for words relative to pseudowords, and for pseudowords relative to words.

Rate effects occurring for both words and pseudowords were identified by looking at effects summed over words and pseudowords (p<0.05 corrected) in regions also showing rate effects for words only and pseudowords only (p<0.001 uncorrected). The interaction between word type and stimulus duration were identified by looking at rate effects occurring for one word type more than the other (e.g. positive linear for words relative to positive linear for pseudowords) (p<0.05 corrected).

Unless otherwise specified, statistical inferences were made at p < 0.05 (corrected), with an extent threshold for each cluster of 10 voxels.

6.3 Results

6.3.1 Effects of Reading relative to Fixation

Activations common to reading words and pseudowords relative to fixation for all stimulus rates, were found in a number of regions including bilateral posterior fusiform, inferior and middle occipital, temporo-occipital, superior parietal, precentral cortices, right frontal operculum, left posterior middle temporal, and bilateral cerebellum (see Table 1). These regions have previously been found to be activated during reading relative to rest in a number of studies (e.g. see Brunswick *et al.*, 1999; Jernigan *et al.*, 1998; Price *et al.*, 1996b).

The left frontal operculum also showed a significant increase in activity for reading pseudowords relative to words at p<0.05 (corrected) (see Table 6.1). This is consistent with previous studies showing the left inferior frontal cortex to be more active for visually presented pseudowords, relative to words (Hagoort *et al.*, 1999; Fiez and Petersen, 1998; Herbster *et al.*, 1997; Rumsey *et al.*, 1997; Price *et al.*, 1996b). The effect of word type on brain activity during reading will be addressed in detail in Chapter 8. The focus of this chapter is on the effect of stimulus rate during reading and its consistency (i) for word and pseudoword stimuli and (ii) across scanning modalities.

These results were replicated in all individual subjects when the statistical threshold was lowered to p<0.01 (uncorrected), apart from two subjects who showed no activity in the left cerebellum and the right anterior middle temporal cortex respectively. This suggests that the detected activations at group level were fairly consistent across subjects.

Anatomical location	Activations for Reading>Fixation								
		<u> </u>	/ords & Pseudo	<u>Words</u>	<u>Pseudo</u>	Difference			
Occipital									
L. posterior fusiform*	-34 -92	-10	9.2	8.7	9.0	n.s			
R. posterior fusiform*	16 -96	-6	9.0	8.6	8.7	n.s.			
L. temporo-occipital*	-46 -68 -	-14	6.0	5.2	6.3	n.s.			
R. temporo-occipital*	46 -60 -	-16	7.6	6.4	7.3	n.s.			
L. inferior and middle occipital*	-44 -76	2	8.0	7.6	7.5	n.s.			
R. inferior and middle occipital*	20 - 100	8	9.3	9.0	9.1	n.s.			
Parietal L. superior parietal R. superior parietal	-28 -66 32 -58	58 58	6.7 7.8	6.4 6.6	7.3 7.5	n.s. n.s.			
Temporal	50 14	,		. 1					
L. anterior middle temporal	-58 -14	-6	4.7	3.1	5.4	3.7 (p>w)			
R. anterior middle temporal	50 -28	-0	4.4	4.9	3.0	n.s.			
L. posterior middle temporal	-52 -56	4	6.7	6.1	5.7	n.s.			
Frontal									
Medial SMA	485	56	6.0	4.5	5.1	n.s.			
L. precentral*	-50 -6 4	46	7.3	4.9	7.2	n.s.			
R. precentral*	52 -12 4	42	6.4	5.0	5.3	n.s.			
L. frontal operculum	-42 4	28	7.2	4.0	7.5	6.3 (p>w)			
R. frontal operculum	40 10 2	28	7.0	5.3	6.5	n.s.			
*									
L. Cerebellum	-46 -64 -	-28	7.7	7.4	7.4	n.s.			
R. Cerebellum	40 -66 -	22	7.9	7.7	7.6	n.s.			

Table 6.1 Regions that showed increased activation during reading relative to fixation for reading both words and pseudowords, words only and pseudowords only. Z-scores which reached significance at p<0.05 (corrected) are reported in bold. Regions also showing any effect of rate (see Table 2) are marked with an asterisk (*). Difference refers to the contrast words vs. pseudowords. Here n.s. = not significant at p<0.001 (uncorrected).

6.3.2 The Effects of Rate Common to Words and Pseudowords

Positive Linear Effects

As seen in the PET study, positive linear effects of rate irrespective of word type were found in visual areas (bilateral posterior fusiform, bilateral temporo-occipital, and bilateral inferior and middle occipital gyri), right superior temporal gyrus and bilateral precentral gyri. All these areas showed effects significant at a corrected level (Z-score > 4.9), apart from the left precentral gyrus (Z-score = 4.2) which was an *a priori* region of interest on the basis of the previous PET study (Price *et al.*, 1996a). At an uncorrected level (p<0.001), a positive linear trend was also detected in the left posterior middle temporal cortex (Z-score = 3.1). These results were replicated in all individual subjects when the statistical threshold was lowered to p<0.01 (uncorrected), apart from one subject who showed no effects in the left posterior middle temporal cortex. This suggests that the effects of rate detected at group level were fairly consistent across subjects.

In addition to the above findings, Price *et al.* (1996a) reported positive linear effects of rate (and an effect of reading relative to rest) in medial cerebellum (Z-score = 7.4), bilateral thalamus (Z-score = 4.7), and medial SMA (Z-score = 4.2). In the present study, only the SMA showed a positive linear trend (Z-score = 1.9) when lowering the statistical threshold to 0.05 uncorrected for multiple comparisons. Nevertheless, the SMA and cerebellum were activated for reading words and pseudowords relative to fixation (p<0.05 corrected for multiple comparisons). The bilateral thalamus, on the other hand, did not show any effect of reading even when the

statistical threshold was lowered to 0.05 uncorrected for multiple comparisons. As I will discuss below, these differences with the previous PET study are likely to reflect subtle differences in task.

Anatomical location	Positive Li of Ra Words & P	near Effect ite for seudowords	Positive Line of Ra for Wo	ear Effect te ords	Positive Linear Effect of Rate for Pseudowords		
Occipital	24 02 12		24 02 10		22 02 12		
L. posterior fusiform	-34 -92 -12	7.5 (9.0)	-34 -92 -12	0.0 (8.5)	-32 -92 -12	5.5 (8.0)	
R. posterior fusiform	32 -88 -14	6.6 (8.2)	32 -88 -14	4.6 (7 .6)	28 - 88 - 14	5.1 (7.6)	
L. temporo-occipital	-44 -68 -14	5.5 (8.3)	-38 -64 -20	4.9 (8.1)	-42 -72 -16	5.1 (8.1)	
R. temporo-occipital	48 - 70 - 6	5.3 (7.7)	48 -70 -6	5.0 (7.3)	44 -80 -4	5.0 (5.7)	
L. inferior and middle occipital	-28 -96 8	5.4 (8.4)	-3096 10	4.5 (7.7)	-30 -96 12	5.1 (8.0)	
R. inferior and middle occipital	22 - 98 8	7.7 (9.2)	22 –98 10	5.1 (8.6)	22 -98 8	7.4 (8.0)	
Temporal							
R. superior temporal	64 - 32 18	5.3 (5.1)	66 – 30 18	4.1 (4.3)	66 -32 16	4.3 (3.6)	
Frontal							
L. precentral	-46 -6 56	4.2 (5.9)	-44 -8 58	3.1 (4.0)	-36 -4 56	3.2 (4.1)	
R. precentral	44 2 56	5.0 (6.2)	42 0 58	3.1 (4.2)	44 2 56	4.8 (5.4)	

Table 6.2 Regions showing positive linear dependencies on stimulus rate. Each response profile was independently assessed for words and pseudowords, words only, and pseudowords only. The first Z-score (not in brackets) refers to the rate effect. The second Z-score (in brackets) refers to the effect of reading relative to fixation at 60wpm. In can be seen that the positive effects of rate were associated with increases in activity for reading relative to fixation. Z-score significant at p<0.05 (corrected for multiple comparisons) are reported in bold.

Negative Linear Effects

Negative linear effects of rate irrespective of word type were not found. In contrast, Price *et al.* (1996a) found a negative linear effect of rate during word reading (and increased activity for reading relative to rest) in both left (x= -52 y = -54 z= 14, Z-score = 5.1) and right (x= 52 y= -60 z= 12, Z-score = 4.5) temporo-parietal junctions. In the present study, there was no effect of rate in the left and right temporo-parietal junctions during word reading even when lowering the statistical threshold to 0.05 uncorrected. Furthermore, these areas were not active for reading words relative to fixation at p<0.05 uncorrected. There was, however, a small negative effect of rate for pseudowords in the left temporo-parietal junction (left (x= -54 y = -64 z= 22, Z-score = 2.5). Negative linear effects of rate irrespective of word type were not detected at an individual subject level either.

Positive and Negative Quadratic Effects

Neither the group or individual analyses revealed positive or negative quadratic effects, irrespective of the word type.

6.3.3 Interactions between Stimulus Rate and Word Type

Interactions between stimulus rate and word type were not found (p<0.05 corrected for multiple comparisons). Thus, there were no regions where the effect of stimulus rate was significantly different for words and pseudowords. Interactions were not detected at an individual subject level either.

6.3.4 Summary of Results

The rate-independent effects in the bilateral parietal and frontal regions are consistent with those detected in the previous PET study (Price *et al.*, 1996a). This replication over scanning techniques lends the effects a face validity. The most notable difference between the PET and fMRI findings was a failure of fMRI to detect negative effects of rate in the bilateral temporo-parietal junctions.



Reading-Fixation

Positive Linear Effect of Rate

Figure 6.1 Brain regions that showed increased activity for reading relative to fixation (upper part) and positive linear dependence on stimulus rate (lower part). Statistical display thresholds were p<0.001 (uncorrected) to show effects in the left precentral gyrus (Z-score = 4.5) which represents an *a priori* region of interest on the basis of the previous PET study (Price *et al.*, 1996a).

Chapter 6

6.4 Discussion

The study presented here has identified linear effects of stimulus rate on brain activity as measured by BOLD contrast, during word and pseudoword reading. In the first instance, I discuss the main response profiles. Then, I consider regions that showed increased activation during reading relative to fixation, but no dependence on stimulus rate. Throughout the discussion, I focus on differences between PET and fMRI signals.

6.4.1 Positive Linear Effect of Rate

As stimulus rate increased, a strong positive linear effect on BOLD signal, common to words and pseudowords, was detected in visual areas, right superior temporal gyrus, and bilateral precentral gyrus. These findings replicate the results of the previous PET study (Price *et al.*, 1996a), confirming that activation in regions associated with visual processing and response generation increases with the number of stimuli per unit time. The consistency between fMRI and PET studies suggests that the findings in these regions are not related to acquisition techniques but reflect true neurophysiological effects. The positive linear effect of rate in visual areas can be parsimoniously explained in terms of processing differences elicited by factors intrinsic to the stimulation itself: the higher the stimulus frequency, the more activity is evoked. However, top-down modulation in both striate and extrastriate visual cortex can not be excluded, as suggested by a number of studies (Gandhi *et al.*, 1999; Somers *et al.*, 1997; Shulman *et al.*, 1997; Heinze *et al.*, 1994; Corbetta *et al.*, 1991).

Another area to show a positive effect of rate was the right superior temporal gyrus, where activity was only elicited at 60wpm. This effect can also be seen in the previous PET data when stimulus duration is 150 milliseconds (see Figure 1c in Price et al., 1996a), and may indicate a specific role for the right superior temporal gyrus. Lesion studies indicate that the right perisylvian area is important for affective components of language processing, such as prosody, musical intonation and emotional gesturing (Ross, 1981). However a number of neuroimaging studies have detected responses in the right superior temporal gyrus for reading single words relative to rest (Jernigan et al., 1998; Price et al., 1996a; Price et al., 1994) and repeating relative to rest (McCrory et al., 2000). Other neuroimaging studies have suggested that the right temporal lobe is important for high level integrative processes needed to achieve global coherence during discourse processing (St George et al., 1999). Further investigation is needed to clarify the function of the right superior temporal lobe, and the right hemisphere in general, in reading. The study presented here suggests that manipulating parametric variables, such as stimulus rate, may help differentiate responses to verbal stimuli in left and right hemispheres.

In addition to the above findings, Price *et al.* (1996a) reported positive linear effects in medial SMA, medial cerebellum and bilateral thalamus. In the fMRI study reported here, none of the above regions was significantly affected by rate (although a positive linear trend was detected in the SMA at a lower threshold). These differences are easily explained by a difference in the task employed by the two studies. In the Price *et al.* (1996a) study, subjects "mouthed" (i.e. silently articulated) the response, thus increasing demands on motor areas (SMA, cerebellum, and thalami) whereas in

the present fMRI study, subjects were instructed to reading silently without articulating.

6.4.2 Negative Linear effects of rate

As stimulus rate increased, negative linear effects on BOLD response were not found for reading words and pseudowords. On the other hand, in the previous PET study, a negative linear effect of rate was found in bilateral temporo-parietal junctions (Z-score = 5.1 on the left and 4.5 on the right). In the present study, these areas were not active for reading words relative to fixation and did not show any effect of rate. Thus, the effects of both stimulus rate and reading relative to fixation were different in the subjects scanned in PET (rCBF) and fMRI (BOLD contrast). Some possible explanations are as follows. The bilateral temporo-parietal junctions have been associated with semantic processing by both lesion and neuroimaging studies (Mummery et al., 1999; Demonet et al., 1994; Warrington and McCarthy, 1987). In the PET study, Price et al. (1996a) attributed the negative linear effect of rate to implicit semantic processing of stimuli when the interstimulus interval was increased. One possibility is that the negative linear effect in bilateral temporo-parietal regions may not have been detected by fMRI because the underlying neuronal signal, associated with implicit semantic processing of stimuli, developed slowly relative to the predicted response (see Veltman et al., 2000).

An alternative explanation is that rather than semantic processing being elicited in response to word presentation, semantic processing is curtailed by word presentation. In this context, the actual presentation of each stimulus transiently

136

interrupts the ongoing set-related activity (see Gusnard and Raichle, 2001; Binder *et al.*, 1999) and the reduction in signal is proportional to the number of stimuli presented. It may be that these implicit deactivations were not detected because fMRI is asymmetric in its ability to detect deactivations relative to activations (due to the inherent nonlinearities in BOLD signal). This corresponds to the "haemodynamic rectification" described in Chapter 4. Figure 6.2 shows that the simulated BOLD response increases with positive stimulus amplitude, but when stimulus amplitude is negative some nonlinear behaviour ensues. On the other hand, the estimated rCBF response is linearly dependent on stimulus amplitude. This suggests PET and fMRI may be differentially sensitive to decreases in neuronal activity compared to increases. See Figure 4.4 (in Chapter 4) for the simulated rCBF and fMRI responses as functions of stimulus amplitude across a wider range of resting oxygen extraction fractions.



Figure 6.2 Simulated rCBF (dotted line) and BOLD (solid line) responses as functions of stimulus amplitude, for relatively high resting oxygen extraction fraction. It can be seen that the simulated BOLD response increases when stimulus amplitude increases, but when stimulus amplitude decreases below 0, nonlinearities are expressed. On the other hand, the estimated rCBF response is linearly dependent on stimulus amplitude over positive and negative values. This suggests that fMRI is asymmetric in its ability to detect deactivations relative to activations (due to the inherent nonlinearities in BOLD signal). Evoked responses were estimated by using the input-state-output model described in Chapter 4.

Chapter 6

6.5 Conclusions

In summary, the data show a pleasing consistency between the effects of visual word ate in PET and fMRI. However, further investigation is needed to clarify whether the discrepancy, relating to the negative effects of rate, is due to the different acquisition techniques or other differences between the two studies. The latter include (i) the two experiments being performed on different subjects (mean age was 57 years 7 months in PET *vs.* 24 years in fMRI), (ii) differences in the precise task (silent articulation in PET *vs.* silent reading in fMRI), (iii) different stimulus duration (150 and 1000 milliseconds in PET *vs.* 600 milliseconds in fMRI), and (iv) different sets of stimuli. It is also possible that discrepancy reflects deactivations that are better detected by PET but not by fMRI.

Chapter 7

Comparing event-related and epoch analysis in blocked design fMRI

7.1 Introduction

In Chapter 1 of this thesis, I addressed the notion that both event-related and epoch models can be used to analyse data collected in blocked design fMRI. Furthermore, I noted that the two models may provide differential sensitivity to experimentally-induced effects. This is because (i) epoch models always predict steady-state dynamics whereas an event-related model may predict a periodically modulated response; and (ii) event-related models may reach its peak earlier and return to baseline later than epoch models, resulting in differential response onsets and offsets.

In this chapter, I explore the impact on sensitivity of differential response onsets and offsets in event-related and epoch models, by using data from the blocked design fMRI study reported in Chapter 6. The design comprised 6 experimental conditions, each presented in blocks of 21 seconds alternated with resting periods of 16 seconds. Conventionally, such a design would be analysed using a boxcar regressor convolved with the HRF. However, in the study presented in this chapter, a mixed model, in which trials were modelled as both epochs and single events, was used in order to compare the sensitivity provided by epoch and event-related analyses. Figure 7.1 shows the resulting epoch (solid line) and event-related (broken line) models when the SOA is 3 seconds. It can be seen that the two models predict differential haemodynamic responses, with the epoch model reaching its peak later and returning to baseline sooner than the eventrelated model (Figure 7.1a). However, both models predict steady-state dynamics because data per voxel were only acquired once every 3.15 seconds (the scan acquisition time). For the event-related model to predict a periodically modulated response, the SOA needs to be long relative to the scan acquisition time, see Chapter 1. When the epoch model is orthogonalised with respect to the event-related model, effects which are modelled by the event-related but not by the epoch model are identified. It appears that the event-related model explains changes in activity at the beginning and at the end of the block that are not accounted for by the epoch model (Figure 7.1b).



Figure 7.1 (a) Difference in onset and offset latencies between epoch (red solid line) and eventrelated (blue broken line) models, when the SOA is 3 seconds. It can be seen that, although the area under the curves is the same, the epoch model reaches its peak later and returns to baseline sooner than the event-related model. (b) When the epoch model is orthogonalised with respect to the eventrelated model, effects which are modelled by the event-related but not by the epoch model are identified. It appears that the event-related model explains changes in activity at the beginning and at the end of the block that are not accounted for by the epoch model.

By modelling trials as both epochs and single events, we were able to estimate the variance in the BOLD signal that was explained by either the event-related or epoch model, above and beyond the other. If the observed response conformed to a HRF convolved with a boxcar regressor, then the event-related model would not evidence an advantage over the epoch model. In contrast, if signal changes reflected the additive contribution of the individual evoked responses to each stimulus, then the event-related model would explain changes in activity more accurately than the epoch model. This would be expressed as a significant effect of reading relative to rest that is explained by the event-related but not by the epoch model. Below I show that the event-related model provides greater sensitivity than the epoch model. Furthermore, I show that the advantage of the event-related model results from its early onset rather then its late offset relative to the epoch model.

7.2 Methods

Chapter 6 provides a detailed description of the experimental design used, the acquisition of the data and the preprocessing applied to the images. In brief, 6 right-handed volunteers were presented with blocks of words or pseudowords alternating with blocks of fixation (a cross in the middle of the screen). The variables were: (i) stimulus type (words and pseudowords) and (ii) stimulus tate (20, 40 and 60 words per minute). An anatomical volume image and a total of 366 functional images with BOLD contrast were obtained for each subject using a 2 T Siemens VISION system. Images were realigned, normalised and smoothed (with a 6 mm full width at half maximum gaussian

filter) using SPM99 software. For the purpose of the study presented in this chapter, three statistical analyses were performed.

7.2.1 Analysis 1: Event-Related and Epoch Regressors

A mixed model was constructed in which events were modelled as both epochs and single events. The resulting design comprised (i) 6 "epoch" effects, in which trials were modelled as blocks using a boxcar regressor, and (ii) 6 "event-related" responses, in which the same trials were modelled as single events. Both were convolved with a HRF. This analysis identified effects of reading relative to fixation (using *t*-tests) that were explained by either model with the other acting as a non-orthogonal confound. However, this analysis did not determine whether the better fit of either model lay in the onset of the response (earlier in the event-related model), the offset of the response (later in the event-related model), or both (see Figure 7.1b). Two further analyses were therefore required.

7.2.2 Analysis 2: Event-Related, Epoch and First-stimulus Regressors

A mixed model was constructed in which events were modelled as both epochs and single events, as in Analysis 1. In addition, a regressor representing the first stimulus in each block, was modelled using an individual delta function. The resulting design comprised (i) 6 "epoch" regressors, (ii) 6 "event-related" regressors, in which the same trials were modelled as single events, and (iii) 6 "first stimulus" regressors, in which the first trial in each block was modelled as a single event. By modelling the first stimulus in each block as a single event, most of the difference in latency onset between the
event-related and the epoch model was discounted. Thus, the residual advantage of either model would be explaind by the differential offset of the expected epoch response (i.e. later in the event-related model).

7.2.3 Analysis 3: Event-Related, Epoch and Last-stimulus Regressors

A mixed model was constructed in which events were modelled as both epochs and single events, as in Analysis 1, with an additional regressor modelling the last stimulus in each block. The resulting design comprised (i) 6 "epoch", (ii) 6 "event-related" and (iii) 6 "last stimulus" regressors. In this analysis, any residual advantage of either model would be explained by the differential onset of the predicted epoch response (i.e. earlier in the event-related model).

The data were high-pass filtered using a set of discrete cosine basis functions with a cutoff period of 120 seconds. The temporal autocorrelations in the errors were dealt with by using a AR(1) + white noise model (Friston *et al.*, 2002). Inferences were made at p<0.05 (corrected for multiple comparisons) with an extent threshold of 10 voxels.

7.3. Results

7.3.1 Analysis 1

When events were modelled as both epochs and single events (Analysis 1), an effect of reading relative to fixation for the event-related model was found in a number of critical language regions (see Figure 7.2 and Table 7.1 for details). In these regions, the event-

related model explained changes in BOLD signal that were not accounted for by the epoch model. In contrast, an effect of reading relative to fixation for the epoch model was not detected, even when lowering the statistical threshold to p<0.001 (uncorrected for multiple comparisons). This means that the epoch model did not explain changes in BOLD signal that were not accounted for by the event-related model.

The effect of reading relative to fixation explained by the event-related model was observed for each stimulus rate and word type (p<0.001 uncorrected). However, greater event-related response was found for reading 20 relative to 60 words per minute in left posterior inferior temporal (x = -34 y = -68 z= -16 Z-score = 4.9), right inferior fusifrorm (x = 16 y= -96 z= -10; Z-score = 3.9) and left inferior frontal (x = -40 y= 6 z= 22; Z-score= 3.6) at an uncorrected level (p<0.001). This trend was expected since, as discussed in Chapter 1, the difference in latency between event-related and epoch models is a function of stimulus rate. This difference would disappear in the limit of very fast presentation rates because the epoch and event-related regressors would be the same.



Figure 7.2 Regions that showed increased activity for reading relative to fixation that was explained by the event-related but not by the epoch model (Analysis 1). These include critical language areas such as left mid fusiform, posterior superior temporal and inferior frontal cortex (see Table 7.1).

7.3.2 Analysis 2 & 3

When the first stimulus in each block was added to the model (Analysis 2), the effect of reading relative to fixation for the event-related model only was no longer detected even when lowering the statistical threshold to p<0.001 (uncorrected for multiple comparisons) (see Table 7.1). In other words, the event-related model no longer explained changes in BOLD signal that were not accounted for by the epoch model. In contrast, when the last stimulus in each block was modelled (Analysis 3), the effect of reading relative to fixation for the event-related model only was still significant, with Z-scores slightly higher than in Analysis 1 (see Table 7.1).

			`			
				Analysis 1	Analysis 2	Analysis 3
Occipital				-	-	-
Left Fusiform	-18	-102	2	5.9	n.s.	6.0
Right Fusiform	18	88	-12	7.0	n.s.	7.7
Left Lingual	-8	-96	-2	6.2	n.s.	7.1
Right Middle Occipital	32	-88	-2	5.6	n.s.	6.8
Temporal						
Left Posterior Inferior Temporal	-36	-66	-18	6.5	n.s.	7.3
Right Posterior Inferior Temporal	38	-66	-18	5.6	n.s.	7.1
Left Superior Temporal	54	-50	10	6.5	n.s.	7.0
Right Superior Temporal	62	-48	10	5.8	n.s.	5.6
Right Middle Temporal	64	-16	-10	6.3	n.s.	7.0
Parietal						
Right Angular gyrus	24	-80	42	6.0	n.s.	6.7
Frontal						
Right Inferior Frontal	44	12	28	6.3	n.s.	6.9
Left Inferior Frontal	-36	6	26	6.1	n.s.	6.5

Reading>Fixation (Event-Related Regressor only)

Table 7.1 Effects of reading relative to fixation that were explained by the event-related but not by the epoch model (Analysis 1). These effects were no longer significant (n.s.) when the difference in the onset of the epoch and the event-related model was discounted (Analysis 2). In contrast, they were still significant when the late offset of the event-related model was discounted (Analysis 3).

7.4 Discussion

In the present study, the relative sensitivity of standard epoch and event-related analyses was investigated using a blocked design reading paradigm. In order to avoid anecdotal comparison of the contrast-specific statistics, which may be confounded by differential efficiency of response estimation for epoch and event-related models, a mixed model was used in which trials were characterised as both epochs and single events (Analysis 1). This allowed direct estimation of the additional variance in the BOLD signal that was explained by either model using the extra sum of squares principle. The results show that, in critical language regions, a significant amount of variance in the BOLD signal was accounted for by the event-related regressor that was not modelled by the epoch-related regressor. Although activation in these regions was reliably detected using an epoch model (see Chapter 6; Mechelli *et al.*, 2000), the present study shows that an event-related approach may characterise the observed response with greater finesse, thereby maximising sensitivity.

In order to determine whether the advantage of the event-related model lay in the early onset or the late offset of the modelled response, the first and the last stimuli in each epoch were also modelled (Analysis 2 & 3 respectively). When the early onset of the event-related model was discounted (Analysis 2), the event-related model no longer showed an advantage. In contrast, when the late offset of the event-related model was discounted (Analysis 3), the event-related model retained its ability to explain changes in BOLD signal that were not accounted for by the epoch model. This indicates that the advantage of the event-related model lies in its early onset rather than its late offset. One possibility for this asymmetry is that adaptation renders the stationary linear model for

evoked responses less appropriate at the end of a block. It should be noted that the advantage of the event related model was still evident when temporal derivatives (which model slight onset and offset differences) were included in the epoch model. This suggests that the advantage of the event-related model was not simply due to a difference in latency but also to a difference in the shape of the haemodynamic response.

Although the shape of the overall haemodynamic response may be affected by nonlinearities in the BOLD signal, especially for short SOAs or long epoch lengths (see Chapter 4), it is important to note that nonlinearities in the amplitude of the haemodynamic response as a function of stimulus rate cannot explain the advantage of event-related over epoch-related analyses. This is because each stimulus rate was modelled separately, allowing for any nonlinear relationship between rate of presentation and response (see Friston *et al.*, 1998 for a full discussion of this in relation to generalized convolution models).

7.5 Conclusion

The use of a mixed model, in which trials were modelled as both single events and epochs, allowed the estimation of the variance in the BOLD signal that was explained by either the event-related or epoch regressors having discounted the effect of the other. The results indicate that, in a number of critical language regions, an event-related model may characterise the form (particularly the onset) of the observed haemodynamic response better than an epoch-related model, even in the context of blocked design fMRI. Better model fit will reduce error variance thereby increasing the sensitivity of the analysis.

Chapter 8

Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies and limitations

8.1 Introduction

This chapter is concerned with why functional imaging studies in different groups of subjects sometimes result in very inconsistent patterns of activation. The paradigm that is explored is one where reading words is contrasted to reading pseudowords – a paradigm which is often chosen to test different cognitive models of reading. Below I briefly describe the models that have motivated a number of functional imaging studies. Then I discuss the ambiguous results and suggest hypotheses to explain the consistencies and inconsistencies in the literature. Finally I present data from an fMRI study that investigates the effect of word type and its interaction with stimulus duration.

8.1.1 Dual Route and Connectionist Models of Reading

Reports in the neuropsychological literature have shown that reading pseudowords can either be more impaired (phonological dyslexia) or preserved (surface dyslexia) relative to reading words with irregular spelling to sound correspondence (e.g. yacht). This observation led to theories that there might be separate neural processes involved in lexical and sublexical reading (e.g. see Marshall and Newcombe, 1973). Although both lexical and sublexical processes may be engaged irrespective of the type of letter string, successful pseudoword reading is reliant on sublexical processes whereas familiar regularly spelled words can be read both lexically and sublexically. The alternative perspective, developed on the basis of computational modeling (Seidenberg and McClelland, 1989; McClelland and Rumelhart, 1981) makes no distinction between lexical and sublexical processes. Rather, the effect of word type results from changes in the connection strength between orthography, phonology and semantics, with more semantic mediation for words than pseudowords. Figure 8.1 illustrates the dual route and connectionist models.



Figure 8.1 Dual route and connectionist models of reading. Solid lines pertain to both models, whereas dotted lines distinguish between distinct routes in the dual route model but not in the connectionist model. The dual route model places greatest emphasis on the distinction between the lexical and sublexical direct mappings from orthography to phonology. In addition a third indirect route is also available via semantics. In contrast, the connectionist model does not distinguish between lexical and sublexical routes. Here the emphasis is on the connection strengths between orthography, phonology and semantics. Both models postulate that phonological retrieval during reading can either occur via direct links between orthography and phonology or via indirect links that map orthography to semantics and semantics to phonology.

8.1.2 Functional Imaging Studies of Word and Pseudoword Reading

Although a number of functional neuroimaging studies have attempted to test alternative models of reading using words and pseudowords (Fiez *et al.*, 1999; Fiez *et al.*, 1998; Rumsey *et al.*, 1997; Herbster *et al.*, 1997), these models do not make clear cut predictions at the neural level. For example, a double dissociation in the activation patterns for reading words and pseudowords could either reflect (i) differential demands on lexical and sublexical routes; or (ii) the degree to which semantics and phonology are activated for words and pseudowords respectively (see Figure 8.1). Interpretation is therefore reliant on a prior understanding of which brain areas are involved in semantic, phonological, lexical and sublexical word processing.

Furthermore, the current neuroimaging literature has not revealed consistent differences between word and pseudoword reading (see Table 8.1 for summary). The most consistent finding is that reading pseudowords relative to words increases activity in a number of language processing areas including the left inferior frontal gyrus (Xu *et al.*, 2001; Fiez *et al.*, 1999; Brunswick *et al.*, 1999; Hagoort *et al.*, 1999; Herbster *et al.*, 1997) and the inferior temporal gyrus (Xu *et al.*, 2001; Brunswick *et al.*, 1999; Price *et al.*, 1996b) but there have been no replications across studies when reading words is contrasted to reading pseudowords.

Word type effects	Price	Herbster	Rumsey	Brunswick	Hagoort	Fiez	Tagamets	Xu
	<i>et al.</i> 1996b	et al. 1997	et al. 1997	et al. 1999	et al. 1999	et al. 1999	et al. 2000	<i>et al.</i> 2001
Words>Pseudowords								
Occipital Left anterior fusiform Left lingual		-36 -30 -24			-18 - 48 0			
Temporal Left middle temporal Right middle temporal Right superior temporal					-50 -34 -12 54 -47 7 61 -36 18			
Frontal Left orbital frontal Left middle frontal Bilateral precentral Bilateral SMA	-22 24 -8 -26 46 28			58 0 44	-3 -19 51			
Other Left parahippocampal Right posterior cingulate Right thalamus			-26 -42 8	14-10 14	11 -20 38			
Pseudowords>Words								
Occipital Left middle fusiform Left middle occipital Right poster. Fusiform Right middle fusiform Right superior occipital Right middle occipital	24 -94 -4				-34 -55 -11 -27 -86 15 44 -48 -10 24 -67 29 25 -83 10	-43 -45 -8	44 -92 -8 44 -64 -28	
Temporal Left medial temporal Left inferior temporal Left middle temporal Left temporo-parietal Right inferior temporal Right middle temporal	-20 -16 -12 -48 -62 -4 -24 -42 28 50 -50 -12 54 -44 4			-42 -54 -20	-5762			-46 -66 -10 32 -76 -14
Parietal Left supramarginal								-38 -48 50
Frontal Left superior frontal Left middle frontal Left inferior frontal Left insula		-44 4 16	-18 20 -4	-48 6 26	-16 -2 46 -46 17 -8	-35 15 6	-36 48 -2	-52 10 12
Other Right cerebellum					12 -43 -16			10 -72 -46

Table 8.1 Summary of the current neuroimaging literature on word and pseudoword reading. All studies used Positron Emission Tomography apart from Tagamets *et al.* (2000) which used fMRI. Statistical thresholds are *not* corrected for the number of comparisons being made. Only those activations that survive a threshold of p<0.001 (uncorrected) are reported.

Inconsistencies may be due to confounds from non-lexical aspects of the stimuli (e.g. differences in bigram frequency, visual input or number of syllables) or they could reflect false positive results due to the low statistical thresholds applied, which were not corrected for the number of comparisons being made. Alternatively, they may reflect (i) true effects that are inconsistent across subjects (different subjects may use different reading strategies); (ii) subtle effects which are not replicated across studies because of insufficient statistical power; or (iii) effects dependent on experimental parameters such as stimulus rate or duration (e.g. greater word type effects for longer stimulus durations).

In this chapter, an fMRI study is presented which compared word and pseudoword reading in 20 subjects. The aim of the study was to (i) investigate consistent and inconsistent differences between word and pseudoword reading, and (ii) distinguish between effects associated with increases and decreases in activity relative to a neutral baseline (fixation). In addition, the study aimed at (iii) investigating whether a significant interaction between word type and stimulus duration could account for some of the inconsistencies in the functional imaging literature in 13 subjects.

In contrast with previous studies of word and pseudoword reading, the data were analysed in a random effect fashion, which allowed inferences to be drawn that pertained to the population from which subjects came as opposed to the particular subjects studied (Friston *et al.*, 1999). When significant effects were not detected, a single subject approach was adopted that enabled the evaluation of whether null results at a group level were due to either a lack of effect in most subjects or significant but inconsistent effects across subjects.

8.2 Methods

8.2.1 Subjects

Informed consent was obtained from 20 right handed volunteers (9 males), aged between 20 and 38 (with a mean age of 26), with English as their first language.

8.2.2 Design and Task

Data were acquired in two different experiments (7 and 13 subjects). Each alternated blocks of words and pseudowords with fixation (to a cross in the middle of the screen). Subjects were instructed to read the words or pseudowords silently as soon as they appeared on the screen. This silent reading paradigm was chosen in order to minimize task-induced effects and avoid the activation of temporal regions caused by subjects processing their own voice (Price *et al.*, 1996b). In Experiment 1 (7 subjects), stimulus duration was constant (600 msec) and stimulus rate varied (20, 40 or 60 words per minute). This experiment included data from the 6 subjects reported in Chapters 6 and 7 and an additional subject performing the same paradigm. In Experiment 2 (13 new subjects), stimulus rate was held constant (40 words per minute) and stimulus duration was manipulated (200, 600 and 1000 msec).

Stimuli were presented in blocks with each of the 6 experimental conditions repeated 5 times in a counter-balanced order across subjects. Each block lasted 21 seconds and was followed by 16 seconds fixation. An eye-tracker was used to monitor the eye-movements of the subjects, in order to ensure that they kept their eyes open and attended to the stimuli.

8.2.3 Stimuli

Stimuli were composed of 4, 5 or 6 letters. Words were matched for frequency (Kucera and Francis, 1967), length and number of syllables between blocks and the grapheme-phoneme relationships were regular (see Section 6.2.3 of Chapter 6 for details).

8.2.4 Scanning Technique

A 2 T Siemens VISION system was used to acquire both T_1 anatomical volume images and T_2 *-weighted echoplanar images with BOLD contrast (see Chapter 2). For each subject involved in either experiment, 366 volume images were taken continuously with an effective repetition time of 3.15 s/volume, the first six (dummy) volumes being discarded to allow for T_1 equilibration effects. Stimulus presentation was arranged so that every 90 msec of peristimulus time was sampled equally over the session (see Chapter 1).

8.2.5 Data Analysis

Data were analysed with SPM99 software using the procedures described in Chapter 3. After pre-processing, a categorical analysis was performed with each of the six conditions (stimulus type x rate or duration) modelled separately in reference to a boxcar waveform convolved with a synthetic HRF. This identified the effects of (i) reading relative to fixation common to words and pseudowords and (ii) word type specifically for each stimulus rate or duration in 20 subjects. In addition, a parametric analysis was performed in which the effects of stimulus duration were modelled as a polynomial expansion of duration up to second order (Büchel *et al.*, 1996). This allowed the nonlinear relationship between neuronal activity and stimulus duration to be investigated in the 13 subjects who performed the duration study. (i) Positive linear dependencies, (ii) negative linear dependencies, (iii) positive quadratic (U-shaped) dependencies; and (iv) negative quadratic dependencies (inverted U-shaped) were identified for both words and pseudowords, for words relative to pseudowords and for pseudowords relative to words independently. The effects of stimulus rate are not addressed here because they have been examined in Chapter 6.

Prior to parameter estimation, a set of discrete cosine basis functions with a cutoff period of 156 seconds was used to high-pass filter the data. Differences in global flow within subjects were removed using grand mean scaling as implemented in SPM99 software. However, given the possibility that the local regional changes might confound the estimate of global signal and lead to "artifactual" decreases, the statistical analyses were replicated without global normalisation. The decreases did not appear to be artefacts and the results reported below refer to the statistical analysis performed using grand mean scaling.

The contrast images from each of the analyses at an individual subject level were entered into one-sample *t*-tests to permit inferences about condition effects across subjects (i.e. a random effect analysis). Unless otherwise indicated, regions that showed significant effects at p<0.05 (corrected for multiple comparisons) with an extent threshold for each cluster of 15 voxels at p<0.001 (uncorrected) are reported.

8.3 Results

8.3.1 Common Effects for Word and Pseudoword Reading

Reading pseudowords and words increased activity relative to fixation (p<0.05 corrected for multiple comparisons) in the left cerebellum and in left posterior inferior temporal, left inferior frontal, and bilateral occipital cortices. Trends (p<0.001 uncorrected) were also found in the left posterior superior temporal and right inferior frontal cortices, bilateral angular gyri and right cerebellum, see Table 8.2 for details. These results are consistent with previous studies on reading (see Price, 2000 for a review), confirming that subjects were engaged in the task.

,	Words & Pseudo >	Fixation	Pseudowords>Fixation	Words>Fixation
Occipital				
L. fusiform	-36 -74 -24	7.0	4.9	4.6
R. fusiform	18 - 86 - 12	5.3	6.2	5.8
Temporal				
L. posterior inferior tempora	1 -40 -62 -16	5.7	6.0	5.1
L. posterior superior tempora	al -56 –50 6	4.1	3.2	3.1
Parietal				
L. angular gyrus	-28 -56 50	4.8	4.9	4.2
R. angular gyrus	34 -54 52	4.8	4.5	4.5
Frontal				
Left inferior frontal	-42 16 32	5.1	3.2	3.1
Right inferior frontal	48 12 30	4.6	3.9	3.6
Other				
Left cerebellum	-36 -52 -26	5.2	4.9	5.4
Right cerebellum	32 -68 -24	4.7	5.3	3.3

Reading > Fixation common to Words & Pseudowords

Table 8.2 Regions that showed an effect of reading relative to fixation common to words and pseudowords. Z-scores significant at p<0.05 (corrected for multiple comaprisons) are reported in bold.

8.3.2 Word type Effects

Pseudowords>Words

Reading pseudowords increased activity relative to words (p<0.05 corrected for multiple comparisons) in the left frontal operculum, left inferior temporal gyrus and right cerebellum, as shown in Figure 8.2. These effects were associated with increases in activity for both words and pseudowords relative to fixation, see upper part of Table 8.3 for details.

Words>Pseudowords

There were no areas where reading words increased activity relative to both pseudowords and fixation (p>0.05 corrected for multiple comparisons). However, the right superior frontal sulcus, right supramarginal, right postcentral and left middle occipital gyrus were more active for words relative to pseudowords. These effects were not associated with an increase for reading words but a decrease for reading pseudowords relative to fixation, see lower part of Table 8.3 for details.

To explore the possibility that different subjects used different reading strategies, effects of reading words relative to pseudowords were addressed at an individual subject level at p<0.05 (corrected for multiple comparisons) with an extent threshold of 15 voxels. Significant effects of reading words relative to pseudowords and fixation were observed in 7 out of 20 subjects (see Table 8.4 for details). However, none of these effects was replicated in any of the other subjects.



Figure 8.2 Brain regions that showed more activity for reading pseudowords than words (p < 0.05 corrected for multiple comparisons), see Table 8.3 for details.

Pseudowords >Words & Fixation

	Pseudowords>Words	Pseudowords>Fixation	Words>Fixartion
Temporal L. posterior inferior temporal	-44 -64 -16 5.3	5.6 (-46 -64 -14 6.8)	5.0 (-46 -66 -12 6.1)
Frontal Left inferior frontal (pars opercularis)	-48 8 22 5.0	4.7 (-44 8 24 5.1)	1.1 (-44 20 20 4.1)
Other Right cerebellum	32 - 70 - 26 5.1	4.7 (30 -68 -26 5.3)	3.0 (38 -66 -24 4.1)

Words > Pseudowords

	Words>Pseudowords	Fixation>Pseudowords	Fixation>Words
Occipital Left middle occipital	-36 -76 28 5.1	4.2 (-42 -78 28 5.7)	0.7
Parietal Right supramarginal	62 - 28 32 5.5	5.0 (64 - 28 34 5.3)	3.5
Frontal Righ postcentral gyrus	56 -18 24 5.3	4.4	1.5
Right superior frontal sulcus	32 20 48 5.0	4.1 (32 18 46 5.0)	1.2

Table 8.3 Regions that showed word type effects in the present study. The effects of pseudowords relative to words were associated with *increases* in activity for both words and pseudowords relative to fixation. On the other hand, the effects of words relative to pseudowords were due to *decreases* in activity for reading pseudowords relative to fixation. Z-scores significant at p<0.05 (corrected for multiple comparisons) are reported in bold.

			W	ords>Pseudo	Words>Fixation
Subejct 1	Left Posterior Cingulate	-10 -46	26	6.0	3.1
	Right Cerebellum	18 -78	-22	8.8	4.3
Subject 4	Left Sup Temp. Sulcus	-38 -58	28	7.8	5.1
Subject 5	Right Middle Frontal	22 62	14	7.3	5.6
-	Right Circular Insular Sulcus	46 -10	20	6.8	4.3
	Left Anterior Cingulate sulcus	-12 56	2	8.4	5.2
Subject 11	Right Anterior Middle Temporal	58 0	-20	6.1	4.9
Subject 15	Medial Paracentral lobule	2 -38	64	5.8	4.1
-	Right Superior Temporal Sulcus	58 – 58	18	7.4	5.7
Subject 16	Right Angular	48 - 48	34	7.2	4.6
Subject 18	Right Posterior Superior Temporal	60 -42	14	5.6	3.2
	Right Posterior Middle Temporal	56 -40	2	7.0	4.1

Words > Pseudowords & Fixation

Table 8.4 Subject-specific effects of reading words relative to pseudowords associated with an increase in activity for reading words relative to fixation. Z-scores significant at p<0.05 (corrected for multiple comparisons) are reported in bold. Although some effects of words relative to pseudowords and fixation were highly significant, there were no replications across subjects.

8.3.3 Duration Effects

Positive Linear Effects

Positive linear effects of stimulus duration were found in bilateral posterior fusiform and middle occipital regions. Positive linear trends (p<0.001 uncorrected) were also found in the left precentral gyrus and cerebellum (see upper part of Table 8.5). These results are consistent with previous PET findings (Price and Friston, 1997).

Negative Linear Effects

Negative effects of stimulus duration were not detected. However, a number of effects significant at p < 0.001 (uncorrected) were found in right-lateralized regions (see lower part of Table 8.5). These largely overlap with the areas that showed negative effects of stimulus duration at p < 0.001 (uncorrected) in Price and Friston (1997).

Quadratic Effects

Quadratic effects, either positive or negative, were not detected even when lowering the statistical threshold to p < 0.001 (uncorrected). The nonlinear dependencies found in Price and Friston (1997) using PET and fixed effect models were therefore not replicated with fMRI when inferences pertained to the population from which subjects came as opposed to the particular subjects studied.

	cts of Duration	
	Positive Linear	Reading 1000 msec > Fixation
Occipital		
Left posterior fusiform	-14 -94 -14	5.3 5.0
	-34 -80 -12	5.2 4.8
Right posterior fusiform	16 -90 -10	5.1 5.0
Left middle occipital	-32 -86 2	5.2 3.8
Right middle occipital	26 -94 10	5.0 4.7
Frontal		
Left precentral	-56 6 38	4.2 3.5
Other		
Left cerebellum	-34 -50 -24	4.2 4.4

Negative Linear Effects of Duration (Trends)

Negative Linear	Reading 200 msec > Fixation
62 -54 0 4.2	2.5
48 -26 -16 4.6	n.s.
54 -56 36 4.5	n.s.
36 6 60 4.0	2.5
44 40 22 3.9	2.4
6 -28 46 3.6	n.s.
32 -4 -4 3.9	2.4
22 -46 -30 4.6	n.s.
-32 -74 -30 4.2	n.s.
	Negative Linear $62 -54 \ 0 \ 4.2$ $48 -26 -16 \ 4.6$ $54 -56 \ 36 \ 4.5$ $36 \ 6 \ 60 \ 4.0$ $44 \ 40 \ 22 \ 3.9$ $6 -28 \ 46 \ 3.6$ $32 \ -4 \ -4 \ 3.9$ $22 \ -46 \ -30 \ 4.6$ $-32 \ -74 \ -30 \ 4.2$

Table 8.5 Regions showing either positive or negative linear effects with stimulus duration. Z-scores significant at p<0.05 (corrected for multiple comparisons) are reported in bold. Here n.s. = not significant at p<0.01 (uncorrected).

8.4 Discussion

The main aim of the present investigation was to establish whether there were any reliable and consistent differences in the pattern of neural activation evoked by word and pseudoword reading. In previous studies, the data were analyzed using conventional fixed effect statistical models, that test for average activation across subjects without segregating "within" and "between" subject variance. This means that significant results at a group level can be found when there are large effects in one or a subset of subjects and no effects in the majority of subjects. Fixed effect analyses allow one to draw inferences pertaining to the particular group of subjects studied but not to the population from which those subjects came. In contrast, the present investigation used data from a larger number of subjects and a random effect statistical model that is based on between subject variance thereby allowing inferences to be drawn at the population level (Friston et al., 1999). Furthermore, in the present study only effects of word type that reach levels of significance corrected for the number of comparisons being made are reported and discussed. This minimizes false positive results, which may be another cause of inconsistencies in the previous functional neuroimaging literature. Below, I discuss the main findings of the present study and their implications for cognitive models of reading.

8.4.1 Main findings

Reading pseudowords relative to words and fixation increased activation in the left frontal operculum, left posterior inferior temporal gyrus (LPIT) and the right cerebellum. In addition, an effect of reading words relative to pseudowords was found in

a dorsal region of the left middle occipital cortex and in right supramarginal, postcentral and superior frontal areas. Critically, the word effects were due to decreased activity for reading pseudowords relative to fixation rather than increased activity for reading words relative to fixation. This raises the possibility that increases for words relative to pseudowords reported by previous studies were due to decreases for pseudowords.

Word type effects due to an increase in activity for reading words relative to pseudowords and fixation were only observed at an individual level. This suggests that inter-subject variability may account for some of the inconsistencies in the neuroimaging literature. It should be noted that a number of subject-specific effects of word type were highly significant with Z-scores as high as 8.8. Such effects may have an impact on the results at a group level in the context of a fixed effect analysis, especially in studies that involve a relatively small number of subjects (i.e. up to 8). In contrast, the random effect analysis used here ensured that significant results were not due to large effects in one or a few subjects only.

The present study also investiagated whether the effects of word type are modulated by stimulus duration. Although significant main effects of stimulus duration were found that were broadly consistent with those reported in Price and Friston (1997), there were no significant interactions between stimulus duration and word type. Furthermore, there was no evidence for an interaction between stimulus rate and word type (Chapter 6). Thus, the present data do not support the hypothesis that the effect of word type is modulated by experimental parameters such as stimulus rate and duration.

8.4.2 Implications for cognitive models of reading

The results of this study, and previous studies of word and pseudoword reading, have not yet provided evidence for one cognitive model of reading over the others. This is because, as indicated in the Introduction, these models do not make clear cut predictions at the neural level. For example, a double dissociation in the activation patterns for reading words and pseudowords could either reflect (i) differential demands on lexical and sublexical routes; or (ii) the degree to which semantics and phonology are activated for words and pseudowords respectively. Furthermore, since pseudowords and words differ in familiarity as well as lexicality, interpreting word type effects requires reference to other studies that have manipulated lexicality or familiarity independently of word type. Here I interpret the effects seen for pseudowords relative to words in reference to the results of other neuroimaging studies that have seen activation in the same areas.

The most robust effects for pseudowords relative to words and fixation were observed in three areas (the left frontal operculum, the LPIT and the right cerebellum). Similar effects for pseudowords relative to words have been reported in the left frontal operculum by Xu *et al.* (2001), Brunswick *et al.* (1999), Hagoort *et al.* (1999), Feiz *et al.*, (1999) and Herbster *et al.* (1997); in LPIT by Xu *et al.* (2001), Brunswick *et al.* (1999) and Price *et al.* (1996); and in both LPIT and the frontal operculum¹ by Paulesu *et al.* (2000). All three areas are also activated during lexical retrieval tasks irrespective of stimulus modality and even in the absence of orthographic input (see Price 2000 for a review). For instance, in the visual domain, they are activated for picture, letter and

¹ Paulesu *et al.* (2000) is not reported in Table 8.1, because the effects did not survive p < 0.001 uncorrected.

colour naming (Price and Friston, 1997); in the auditory domain, they are activated for repetition (Noppeney and Price, 2002) and word generation (Frith *et al.*, 1991); and in the tactile domain, they are activated when blind subjects read braille (Büchel *et al.*, 1998). Thus, increased activation for pseudoword relative to word reading corresponds to that found for lexical retrieval. Previous studies have suggested that pseudowords may increase the demand on lexical processing because there is no semantic representation available (Fiez *et al.*, 1999; Price *et al.*, 1996). However, this conclusion does not allow one to determine whether there are separate neural systems involved in lexical and sublexical orthographic to phonological mapping. For instance, the common activation pattern observed for words and pseudowords might include a sublexical processing system that is equally activated by words and pseudowords (e.g. in the left posterior superior temporal cortex). It is also possible that sublexical and lexical processes engage different neuronal populations within the same anatomical region.

While the effects in the LPIT and the frontal operculum are relatively consistent with the literature, it can be noted that greater activation in the right cerebellum for pseudowords relative to words was not reported in any of the studies listed in Table 8.1. One possible explanation is that, in this region, the effect of word type is subtle but highly consistent across subjects. This would result in significant activation when the effect size is compared against inter-subject variability (i.e. in a random effect analysis) but not when the effect size is averaged across subjects and compared against the within subject variability (Friston *et al.*, 1999b).

With respect to the decreases in activity observed for reading pseudowords relative to fixation, localized task-induced decreases in activity are a relatively frequent

finding in functional neuroimaging (see Mazoyer et al., 2001; Binder et al., 1999; Shulman et al.,1997) but remain poorly understood. Binder et al. (1999) suggest that rest/fixation is a semantic condition, therefore decreased activity for pseudoword but not word reading may reflect an interruption of ongoing semantic activity. However, the areas that showed decreased activity for pseudoword reading (right superior frontal sulcus, right supramarginal, right postcentral and left middle occipital gyrus) do not correspond to those associated with semantic processing. An alternative perspective (see Gusnard and Raichle, 2001) is that the decreased activity observed during task performance correspond to neuronal deactivations. Such deactivations would occur in areas that are not involved in task performance to facilitate the processing of task relevant information. Within this framework, deactivations for pseudowords relative to words might reflect the relative demands on the reading system. If this is indeed the case, the data presented here illustrate that fMRI can be sufficiently sensitive to neuronal deactivations in spite of the BOLD rectification effects described in Chapter 4.

Finally, the lack of consistent activations for words relative to pseudowords and fixation needs to be addressed. Here there are a number of possible explanations. For instance, words may engage the same neuronal components as pseudowords with word specific effects emerging only in the strength of functional connections. This may not be revealed by classical subtraction paradigms but requires tests of functional or effective connectivity such as the application of structural equation modeling (see Chapter 9; Horwitz *et al.*, 2000; Horwitz *et al.*, 1998; Büchel and Frison, 1997; McIntosh and Gonzalez-Lima, 1994). Alternatively, if word and pseudoword processing is distinguished by differences in the timing of activation, this will not be detected when

activation is summed over time. For example, onset or peak activation at the same location may occur at different times for words and pseudowords (Simos *et al.*, 2002; Henson *et al.*, 2002) or one word type might evoke a transient response while the other evoked more sustained activity at a lower amplitude. In summary, there are many different ways that word specific effects could be expressed at the neural level but the imaging techniques currently available may not be sufficiently sensitive to detect them.

8.5 Conclusions

In this chapter, I have identified consistent and inconsistent effects of word type across subjects, and distinguished between effects associated with increases and decreases in activity relative to fixation. Reading words and pseudowords activated a common set of areas and, within this system, pseudowords increased activation in areas previously linked to lexical retrieval. In contrast, consistent effects of reading words relative to pseudowords were only found in regions that showed decreases in activity for pseudowords rather than increases for words relative to fixation. Word specific activations relative to fixation were observed at the individual level but with no consistency over subjects.

These results (i) illustrate that pseudowords place increased demands on areas that have previously been linked to lexical retrieval, and (ii) highlight the importance of including one or more baselines to qualify word type effects. Furthermore, (iii) they suggest that inconsistencies observed in the previous literature may result from effects arising from a small number of subjects only.

Chapter 9

Effective Connectivity and Inter-Subject Variability: using a Multi-Subject Network to test Differences and Commonalities

9.1 Introduction

In Chapter 8, I investigated the effect of word type by directly comparing brain activity for reading words and pseudowords. While reading pseudowords relative to words and fixation increased activation in a number of regions consistently over subjects, reading words relative to pseudowords and fixation only showed effects at an individual level. In this chapter, I explore the effect of word type on the functional integration of activity in different brain regions involved in reading. Specifically, I illustrate how differences in effective connectivity among subjects can be addressed explicitly using Structural Equation Modelling. Differences in functional integration across subjects may reflect varied cognitive strategies adopted to perform a task. Alternatively, different patterns of connectivity may reflect alternative solutions the brain can adopt to complete the same strategy.

First, I review briefly how previous neuroimaging studies of functional and effective connectivity have dealt with differences and commonalities among subjects. Second, I describe how the issue of inter-subject variability can be addressed directly and illustrate the approach with empirical data.

9.1.1 Group and Subject-Specific Analysis of Functional Integration

Over the past decade, a number of Positron Emission Tomography and functional Magnetic Resonance Imaging studies have tried to characterise functional connectivity (i.e. the temporal correlations among neurophysiological events in different neural systems) and effective connectivity (the influence that one neural system exerts over another). While some PET studies have estimated connectivity by exploiting subject-tosubject variability (e.g. see Horwitz et al., 1998), most studies have discounted intersubject variability by treating data from different subjects as if they came from the same subject (Iidaka et al., 2001; Bokde et al., 2001; Bullmore et al., 2000; Maguire et al., 2000; Coull et al., 1999). Here estimates of the coupling amongst regions of interest are based on item-to-item or block-to-block variability over subjects. This approach assumes that the pattern of connectivity estimated over subjects is a good approximation to the underlying connectivity in all the subjects studied. Treating data from different subjects in this way allows one to draw inferences about the group of subjects studied under the assumption that variations in connectivity from subject to subject are random, wellbehaved and uninteresting. It also increases sensitivity by augmenting the number of data points per voxel. However, this approach may be problematic when inter-subject variability is pronounced. For example, when there are subjects with opposite patterns of connectivity, the estimated connections may be far from the true values in any of the subjects studied. This was illustrated by Gonçalves et al. (2001), who found that the results of the group analysis did not reflect any of the features seen in individual analyses. In addition, estimating connectivity over subjects does not allow one to relate

clinical or demographic measures (e.g. task performance or age) to indices of functional integration (i.e. correlations among regions) in a subject-specific fashion.

An alternative approach involves analysing the data by performing subjectspecific analyses (Gonçalves *et al.*, 2001; Mechelli *et al.*, 2001; Büchel and Friston, 1998; Büchel and Friston, 1997). This method allows one to estimate connectivity independently for each subject. In addition, it enables one to correlate behavioural measures with indices of functional integration over subjects. However, this approach indicates apparently inconsistent results among subjects (Gonçalves *et al.*, 2001; Mechelli *et al.*, 2001). This raises the question of whether *inconsistent effects* correspond to *significant differences* in connectivity.

In summary, unless the model of the data can accommodate subject-specific variations in connectivity, differences among subjects cannot be evaluated. The aim of the present work is to show how individual differences in connectivity can be addressed directly using Structural Equation Modelling (see Chapter 3).

9.1.2 Using a Multi-Subject Network to test Differences and Commonalities

In the study presented here, a multi-subject network is used which comprises p regions of interest for each of the n subjects studied, resulting in a total of $p \times n$ nodes. Regions pertaining to the same subject are (reciprocally) connected to allow for coupling between different areas of the same brain. In contrast, regions from different subjects are not connected. Differences in functional integration across subjects are tested by comparing a model which allows for inter-subject variability with one that does not. This is implemented by setting up a model in which one or more effects of interest (e.g. the

modulation of coupling by stimulus type) may vary over subjects versus one in which these effects are held to be equivalent in all subjects. The best model is the one that explains most of the observed variance-covariance structure for the least cost, in terms of the number of model parameters. This is analogous to testing for subject by condition interactions in classical ANOVA.

Constructing a network of regions from different subjects may seem counterintuitive but embodies two key advantages. First, it allows one to test whether changes in connectivity differ significantly among subjects. Note that significant differences among subjects may occur not only when subject-specific analyses of effective connectivity give inconsistent inferences, but also when subjects show the same effect to a different extent. Second, constructing a network which comprises p regions from each of the n subjects studied provides one with enough non-redundant elements in the covariance matrix $e = \frac{1}{2}(p \times n)(p \times n + 1)$ to estimate all potential connections among regions pertaining to the same brain $q = n \times (p \times (p-1))$ and all residual variances $p \times n$. In essence, by pretending that all the subjects are scanned at the same time, a multi-subject network can be formed that ensures a high degree of sparsity. Note that $e - q \ge 1$ is a necessary but not sufficient condition for a model to be globally identifiable (this is known as the 't-rule'; see Bollen, 1989 p.93). A model is globally identifiable if it is possible to obtain unique values for each free parameter (e.g. connection strength). Global estimability is not always simple to evaluate in practice, especially when the model contains many free parameters, but can be easily detected using computer programs for SEM such as LISREL (Jöreskog and Sörborn, 1993) or EQS (Bentler, 1992).

In contradistinction, when SEM is used to estimate connectivity averaging over subjects or independently for each subject, the number of parameters which can be estimated $e = \frac{1}{2}p \times (p+1)$ is fewer than the maximum number of connections which can be specified $q = p \times (p-1)$ and the residual variances (p) which need to be computed. Usually, the t-rule is enforced either by constraining the number of specified connections or by assigning *a priori* arbitrary values to some of them. However, assuming that one region does not exert an influence over another may be problematic because primate anatomical data show that each region is reciprocally interconnected to a large number of other regions. In addition, assuming a particular *a priori* value for the coupling between two regions may be difficult when data from previous studies are not available.

The multi-subject network approach used here also differs from previous applications of SEM to functional imaging data in terms of the type of variance that is modelled. Most SEM studies in functional neuroimaging do not explicitly model the influence of exogenous or experimental variables. Rather, they assume that the estimated path coefficients reflect endogenous or intrinsic variance only. However, identical regions in different subjects are likely to be correlated because of the common influence of exogenous variables. If not accommodated, these induced covariances would compromise the goodness of fit of the model. The common influence of the exogenous variables is modelled by connecting a virtual node, whose time course reflects stimulus onsets, with the first sensory or "input" region in all subjects. Further experimental changes in task or cognitive set enter through modulation of the connections. Note that the modelling of both endogenous and exogenous variance is not specific to the use of a multi-subject network but is also possible in the context of individual subject and group analysis. It is an important device because it allows explanatory or design variables to explain the inter-regional covariance induced. This approach is particularly essential in multi-subject networks if the subjects are all exposed to the same paradigm. Below a study is presented, which uses a multi-subject network to investigate inter-subject variability of functional integration during single word reading. Specifically, I tested whether the modulatory effect of word type on the reading-induced coupling was significantly different among subjects.

9.1.3 The Neural Correlates of Single Word Reading

As discussed in Chapter 8, neuropsychological data indicate that there are at least two different cognitive strategies that can be used to read, one that relies on access to lexical semantic information and one that involves direct links between spelling (orthography) and sounds (phonology). Familiar words with unusual spellings (e.g. CHOIR) rely on lexical semantic access whereas unfamiliar "pseudowords" which have no meaning (e.g. CHOIN) rely on sublexical spelling to sound relationships. However, functional imaging studies have not shown a clear double dissociation at a neuroanatomical level for word and pseudoword reading (see Chapter 8 for a review). Studies of effective connectivity may provide an alternative means for dissociating the different reading routes (Horwitz *et al.*, 1998). This approach relies on *a priori* anatomical models. In the study below, I use the anatomical model of reading developed in Price (2000), see Figure 9.1.



Figure 9.1 Anatomical model of reading developed by Price (2000).

The model distinguishes two routes for reading, one that is shared by objects and involves the left posterior inferior temporal cortex (the semantic route) and one that is not activated by object naming and involves the left posterior superior temporal cortex (Wernicke's area). Although both routes may be activated irrespective of whether words or pseudowords are being read, the effective connectivity between the left posterior inferior temporal area and Wernicke's area may differ for word or pseudoword reading. Furthermore, connectivity changes may depend on individual reading strategies. For example, reading words with regular spelling may depend on sublexical processing for some subjects but lexical processing for other subjects. In the last part of the chapter, I discuss the results and further applications of the multi-subject network analysis.

9.2 Methods

A detailed description of the experimental design used, the acquisition of the data and the analysis performed is provided in Chapter 8 (Experiment 2). In brief, 13 right handed volunteers were presented with blocks of words or pseudowords alternating with blocks of fixation (a cross in the middle of the screen). The variables were: (i) stimulus type (words and pseudowords) and (ii) stimulus duration (200, 600 and 1000 msec).

The regions entering into SEM were defined using a conventional statistical parametric mapping analysis. After pre-processing, each of the six experimental conditions was modelled independently. A statistical parametric map of the main effect of reading was generated independently for each subject. In addition, the contrast

images from each subject were entered into one-sample *t*-tests to permit inferences about condition effects across subjects (i.e. a random effects analysis).

9.2.1 Specification of the Multi-Subject Network

On the basis of the anatomical model of reading developed by Price (2000), 4 leftlateralized anatomical regions were selected in the left posterior fusiform, posterior inferior temporal cortex, posterior superior temporal cortex and frontal operculum. As predicted by the model, these areas were activated for reading relative to fixation in 10 out of 13 subjects (p<0.001 uncorrected) and at the group level (p<0.05 corrected for multiple comparisons). The three subjects who showed no activation of the inferior temporal gyrus for reading relative to fixation were excluded from further analysis. Each region (6 mm radius) was selected specifically for each subject, to correspond with the most significant voxel in the categorical comparison. Regional activity was defined as the first eigenvariate of any region as in Büchel and Frison (1997). All time series were adjusted for confounds (e.g. global mean, low frequency components) after applying the general linear model with condition specific predictors (Friston *et al.*, 1995b). Figure 9.2 shows the positions of the selected regions in the 10 subjects.



Figure 9.2 The subject-specific regions that entered into the SEM. These comprised four leftlateralized regions: posterior fusiform (average co-ordinates x=-32 y= -87 z= -13), inferior temporal (x= -48 y= -57 z= -13), superior temporal gyrus (x= -47 y= -51 z= 13) and frontal operculum (x= -45 y=19 z= -2).

A multi-subject structural equation model was constructed which comprised the four regions above and all potential forward and backward connections amongst them within each subject. The resulting multi-subject network contained $10 \times 4 = 40$ nodes and $10 \times 12 = 120$ connections. In addition, a virtual node, whose time course corresponded to stimulus onsets, was connected with the posterior fusiform region in all subjects which resulted in a total of 41 nodes and 130 connections. This allowed the common influence of stimulus presentation on the induced inter-regional covariance to be accounted for in all models tested (see Introduction).

9.2.2 SEM Analysis

Structural equation modelling was performed using the SEM Toolbox of SPM99 (Rowe *et al.*, 2002; Maguire *et al.*, 2000; Coull *et al.*, 1999), which employs an iterative maximum likelihood algorithm (Higham, 1993) to estimate covariances that best predict the observed variance-covariance structure of the empirical data. The maximum likelihood algorithm assumes multi-normality in the data, although the multi-subject network contained variables such as word type that were not normally distributed. However, the maximum likelihood algorithm is robust to normality deviations (see Bollen, 1989 p.126) and the final results should not differ markedly if distribution-free methods were to be used. The residual influences were dealt with by fixing them to an arbitrary value as in Büchel and Friston (1997) in all regions/subjects.

Estimation of the Goodness of Fit of the Multi-Subject Network

First, the goodness of fit of the multi-subject network was estimated by performing a χ^2 test (p<0.05). Here reciprocal connections (e.g. Posterior Fusiform \rightarrow Inferior Temporal

and Inferior Temporal-Posterior Fusiform) were set to have the same value within each subject in order to ensure that the model could be globally identified. In contrast, no connections were constrained to be the same over different subjects. The χ^2 test assessed whether the multi-subject network, which allowed for inter-subject variability, accommodated the variance-covariance structure of the empirical data reasonably (i.e. was not an implausible model). Since the meaning of significant path coefficients in a poorly fitting model is unclear (MacCallum, 1995), the multi-subject network was required to provide a good fit according to the χ^2 index.

SEM of Word Type Effects over Subjects

Second, the impact of word type was tested on each connection independently (e.g. the impact of words vs. pseudowords on the connection Posterior Fusiform \rightarrow Inferior Temporal). Here the connection of interest was allowed to differ from its reciprocal connection. However, the remaining reciprocal connections were set to have the same value within each subject. This allowed the impact of word type on each specific connection to be tested without compromising the global estimability of the model.

The impact of word type on connectivity was modelled using a set of virtual nodes whose time course represented the interaction between source activity and the appropriate design variable, as described in Büchel and Friston (1997). These interaction terms were constructed by multiplying the mean-centred time series of the source region by the design variable (in the analysis a vector that encoded the presentation of pseudowords *vs.* words). The activity in any "target" region could than be characterised
in terms of a linear combination of (i) activity in the source region; (ii) the design variable; and (iii) the interaction term. The influence of the interaction term on the "target" area corresponds to the effect of word type on the coupling between the "source" and "target" regions, as depicted in Figure 9.3. The interaction terms play exactly the same role as psychophysiological interactions (PPIs) in simple regression analyses of inter-regional coupling (Friston *et al.*, 1997).



Figure 9.3 The model for the modulatory effect of word type on connectivity (left) and its implementation in SEM (right). Activity in the "target" region (i.e. frontal operculum) was characterised in terms of a linear combination of (i) activity in the "source" region (i.e. superior temporal); (ii) the design variable (here a vector that encoded the presentation of pseudowords *vs.* words); and (iii) an interaction term constructed by multiplying the mean-centred time series of the source region by the design variable. The influence of the interaction term on the target region corresponds to the effect of word type on the coupling between the source and target regions.

The significance of the modulation of connectivity by word type was assessed, independently for each connection, by comparing a model in which this stimulusspecific effect was set to be *zero* with one in which it was *the same* in all subjects. The best model was identified as the one which explained as much of the observed variancecovariance structure as possible but required the fewest number of model parameters. The significance of the difference between null and alternative models was expressed as the difference in χ^2 goodness of fit with 1 degree of freedom (p<0.05). A total of twelve comparisons (i.e. one for each connection) were performed to investigate the impact of word type on connectivity, which involved 24 models (12 models in which the effect of word type on each connection of interest was set to be the same in all subjects and 12 models in which the same effect was set to be zero in all subjects). The multi-subject network used to investigate the effect of word type on the forward connection between the superior temporal cortex and the frontal operculum is shown graphically in Figure 9.4. Similar networks were used to investigate the effects of word type on the remaining connections. Regional nodes are in white and virtual nodes are rendered in gray.

It should be noted that, although reciprocal connections were constrained to be the same, context or condition-sensitive changes in connectivity were free to be asymmetric. This is because the modulations by word type in each direction were modelled by different virtual nodes. This is important because, although anatomical connections are generally reciprocal, they show asymmetries in their functional expression with backward connections being more modulatory (Friston and Price, 2001).



Figure 9.4 The multi-subject network used to look at the effects of word type on the readinginduced coupling between the Superior Temporal gyrus and the Frontal Operculum (similar networks were used to investigate the effects of word type on the remaining connections). The network was composed of the subject-specific regions in Figure 9.2. A virtual node, encoding stimulus onsets, was connected to the posterior fusiform gyrus in all subjects to account for the common influence of stimulus presentation. Word type dependent changes in connectivity were modelled using a further set of virtual nodes whose time course represented the interaction between source activity and the appropriate design variable. Anatomical nodes are in white and virtual nodes are in gray.

SEM of Inter-Subject Variability

Inter-subject variability of this modulatory effect was assessed independently for each connection by comparing a free model (in which the effect of word type on connectivity was free to vary across subjects) with a restricted model (in which the effects of word type on connectivity were set to be the same in all subjects). Again, the best model was identified as the one which explained as much as possible of the empirical correlations with the fewest number of estimated parameters. In this case, the significance of the difference between models was expressed as the difference in χ^2 goodness of fit with 9 degrees of freedom (i.e. number of subjects minus one). A total of twelve comparisons (i.e. one for each connection) were performed to investigate whether the impact of word type on connectivity varied across subjects. This involved comparing the 12 models in which the effect of word type on each connection of interest was set to be the same in all subjects with 12 models in which the same effect was free to vary across subjects.

9.3 Results

The goodness of fit of the multi-subject network, in which reciprocal connections were set to be the same within subjects but free to vary among subjects, was 795.58. This corresponded to a p-value of 0.12 under an asymptotic χ^2 distribution with 115 effective degrees of freedom (computed after correction for autocorrelation in the fMRI time series; see Worsley and Friston, 1995). The model was therefore not refuted (i.e. it provided a sufficient account of the data). With respect to the modulation of connectivity by word type, a number of functional connections were stronger when reading pseudowords relative to words. These included both forward (Posterior Fusiform—Inferior Temporal, Posterior Fusiform—Frontal Operculum, and Superior Temporal—Frontal Operculum) and backward (Frontal Operculum—Superior Temporal) connections. Significant decreases in connectivity when reading pseudowords relative to words were not found. These findings are summarised in Table 1. The values reported represent the modulation by word type on the coupling between a "source" and a "target" region. A value significantly higher than zero indicates that the coupling between the two regions is stronger during reading pseudowords than words, while a value significantly lower than zero indicates that the coupling is weaker.

All forward effective connections were better characterised by a model in which changes in connectivity for pseudowords relative to words were set to be the same across subjects. In contrast, the backward connection between the Frontal Operculum and the Superior Temporal gyrus showed significant inter-subject variability. In 7 out of 10 subjects, this connection was stronger when reading pseudowords but the size of this effect was highly variable. The remaining subjects showed either a decrease in connectivity for reading pseudowords relative to words or no effect at all. P-values of the difference between the free model (which allowed for inter-subject variability) and the restricted model (in which effects were set be the same across subjects) are reported in the last column of Table 1. Results are represented graphically in Figure 9.5.

	Free model (allowing for inter-subject variability)											Free vs. Restricted
	subj.1	subj.2	subj.3	subj.4	subj.5	subj.6	subj.7	subj.8	subj.9	subj.10	subj.1-10	p-value
Forward connections												
Post. Fus. \rightarrow Inf. Temp.	0.22	0.11	0.34	-0.08	0.19	-0.13	0.08	-0.03	0.33	0.25	0.11*	0.15
Post. Fus. \rightarrow Sup. Temp.	0.24	0.05	0.00	-0.02	0.06	0.00	-0.18	-0.25	-0.15	-0.23	-0.03	0.24
Post. Fus.→Frontal Operc.	0.13	0.05	-0.10	0.08	0.24	-0.05	0.14	0.17	0.15	0.11	0.10*	0.89
Inf. Temp.→Sup. Temp.	0.17	0.12	0.22	0.03	0.27	-0.11	-0.05	-0.06	-0.16	-0.24	0.01	0.19
Inf. Temp. \rightarrow Frontal Operc.	0.08	0.06	-0.06	-0.02	-0.06	0.01	0.08	0.01	-0.09	0.05	0.00	0.99
Sup. Temp.→Frontal Operc	. 0.41	0.09	0.00	0.19	-0.09	0.16	0.24	0.07	0.00	-0.13	0.10*	0.11
Backward connections												
Inf. Temp. \rightarrow Post. Fus.	0.14	0.12	0.18	-0.14	0.06	-0.07	0.08	-0.02	0.28	0.00	0.05	0.61
Sup.Temp.→Post. Fus.	0.15	0.07	0.01	0.01	0.00	-0.01	-0.13	-0.09	-0.08	-0.04	-0.01	0.92
Frontal Operc. \rightarrow Post. Fus.	0.13	-0.05	0.12	0.04	0.06	-0.08	-0.03	0.08	0.11	-0.11	0.02	0.83
Sup.Temp.→Inf.Temp.	0.07	0.06	0.13	0.08	0.19	-0.12	-0.05	-0.01	-0.01	-0.14	0.01	0.67
Front. Operc.→Inf.Temp.	0.06	0.03	0.01	0.00	-0.11	0.05	0.10	0.00	-0.18	0.18	0.01	0.68
Front. Operc.→Sup.Temp.	0.48	0.11	0.00	0.06	-0.06	0.17	0.23	0.05	0.04	-0.21	0.10*	0.01

Effects of Reading Pseudowords vs Words on Connectivity

Table 9.1 The effect of word type on the coupling between a "source" and a "target" region in the free model (which allows for inter-subject variability) and the restricted model (in which effects are constrained to be the same in all subjects). A value higher than zero indicates that the coupling between the two regions is stronger during reading pseudowords than words, while a value lower than zero indicates that the coupling is weaker. The asterisk (*) indicates significant effects of word type on the reading-related covariance, when changes in connectivity are set to be the same in all subjects (p-value<0.05). The last column refers to the p-values of the difference in goodness of fit between free and restricted models. It can be seen that, for the backward connection between the Frontal Operculum and the Superior Temporal gyrus, the free model accounts for the data better than the restricted model (i.e. p<0.05).



Figure 9.5 Graphical representation of the SEM results. Solid black lines indicate connections that were modulated by word type but did not show significant inter-subject variability. Broken black lines indicate connections that were significantly modulated by word type and showed significant inter-subject variability. Broken gray lines indicate connections that showed neither a modulation by word type nor significant inter-subject variability.

9.4 Discussion

In the present study, a multi-subject network was used to investigate inter-subject variability in functional integration in the context of single word and pseudoword reading. Within each subject, each region was reciprocally connected to the remainder by both forward and backward connections. This was motivated by the idea that functional specialisation is not an intrinsic property of any region, but depends on both forward and backward connections (Friston and Price, 2001; McIntosh, 2000; Hebb, 1949). To ensure that the model could be globally identified, reciprocal connections were forced to have the same value. In addition, exogenous variance was modelled to account for the common influence of experimental effects over subjects. These influences were encoded (i) by design variables that entered the "input" regions directly

to model sensory evoked responses and (ii) indirectly through interactions with regional activity to model context-sensitive changes in connection strength.

9.4.1 Goodness of Fit of the Multi-Subject Network

First, I evaluated whether the multi-subject network which allowed for inter-subject variability provided a good explanation for the variance-covariance structure of the empirical data. A critical principle in model specification and evaluation is that, by definition, all of the models are wrong or incomplete to some degree. In other words, models at their best can provide only a close approximation to the observed data, rather than an exact fit. In the case of the present study, the neuronal dynamics underlying the reading of words and pseudowords are likely to involve more regions (e.g. right hemisphere areas) than those included in the model. In addition, such dynamics are likely to be more complex than we can represent using a structural equation model which only embodies first order interactions. The finding that the multi-subject network fitted the observed data means only that the model provides a "plausible" representation of the real neuronal dynamics. This enabled the use of the multi-subject network to investigate the effect of word type on connectivity.

9.4.2 The Effect of Word Type on Effective Connectivity Over Subjects

The results showed that a number of forward and backward connections were modulated by word type. A double dissociation in the pattern of reading activity for words and pseudowords was not revealed because the change in connection strength was in one direction only - stronger for pseudowords. In most cases, only the forward connection was modulated by word type whereas the backward connection was less affected. However, in the case of the posterior superior temporal region and the frontal operculum, both the forward and backward connections were stronger during reading pseudowords than words. These findings illustrate that differences between processing words and pseudowords do not simply lie in the degree of activation in one or more regions of the language system. Rather, such differences can be characterised in terms of context-sensitive interactions among brain areas. Connectivity in both semantic and non-semantic reading routes was stronger for pseudowords. One possibility for this result is that, if functional connectivity is a function of task difficulty (Furey *et al.*, 2000), pseudowords are associated with greater connectivity than words because they are more difficult to read.

9.4.3 Effective Connectivity and Inter-Subject Variability

The main aim of the present study was to characterise inter-subject variability by testing for differences in connectivity among subjects. The backward connection between the frontal operculum and the superior temporal gyrus was modulated by word type differentially across subjects. This illustrates that a pattern of connectivity, which is estimated over subjects, may not be a good approximation of the underlying patterns of connectivity in all of the subjects studied. Rather, in some cases a model which allows for inter-subject variability would be preferred to a model in which one single pattern of connectivity is assumed. The results suggest that the subjects may have adopted different strategies either for reading words or pseudowords, which influenced the backward connections between frontal operculum and posterior superior temporal region. For instance, pseudowords are fairly ambiguous stimuli that may engage a search for (missing) semantic and phonological representations in some subjects but not in others (Fiez and Petersen, 1998; Price *et al.*, 1996b). However, the results should be considered in the context of the experimental paradigm adopted in the present study. Silent reading of 40 single words per minute is a relatively undemanding task which is likely to leave additional attentional resources that may be directed at other cognitive processes. One possibility is that, if subjects performed a more engaging task, functional integration would have been more consistent across subjects. Finally, differences in functional integration may reflect alternative solutions the brain can adopt to complete the same strategy.

9.4.4 Further Applications of the Multi-Subject Network Analysis

I have described how the issue of inter-subject variability in functional integration can be addressed directly and I have illustrated the approach with analyses of experimental data. Another possible application involves comparing changes in connectivity among different groups of subjects. This may be particularly useful in the context of patient studies, in which one may want to compare functional integration between patients and normal subjects (e.g. dyslexic vs. normal) or between different groups of patients (e.g. surface vs. phonological dyslexia). In this case, a multi-subject network can be treated as a multi-group network. If the model, which allows for group differences, accounts for the empirical data better than a restricted model in which the groups are the same, then the two groups express differential connectivity. The crucial advantage of using a multigroup network over performing a group or a single subject analysis is that one can make inferences about differences among groups *in addition to* inferences pertaining to a single group or subject, in the same way classical analysis of variance allows one to test for group x condition interactions. Finally, a multi-subject network can be used to address inter-subject variability in relation to behavioural measures. For instance, one can investigate whether differences in performance correspond to differences in functional integration among groups. This approach enables one to interpret changes in connectivity meaningfully by establishing a relation between changes in connectivity and behaviour.

Chapter 10

General Conclusions

10.1 Summary of the thesis

The work in this thesis investigated a number of factors that affect sensitivity to language activations in fMRI. Here sensitivity is defined as the ability to detect experimentally-induced activations, which is a function of (i) the size of the effect of interest and (ii) the efficiency of the design with which the haemodynamic responses are estimated.

Chapter 4 was concerned with the coupling between neuronal activity and perfusion changes across a range of experimental design parameters including epoch length, SOA and stimulus amplitude. Using biophysical simulations based on a haemodynamic input-state-output model, the optimum parameters for maximising BOLD signal changes *per stimulus* or *event* were estimated. The results showed that, when the SOA is 1 second, the BOLD response decreases with epoch length and levels out around 10 seconds. This means that there is only little gain in limiting the epoch length to 10 seconds instead of 20 or 30 seconds. The results also showed that the BOLD response increases with the SOA until it reaches a maximum around 8 seconds. As the SOA increases further, the BOLD signal decreases slightly until it reaches a plateau around 12 seconds. It was also demonstrated that fMRI appears to be asymmetric in its ability to detect deactivations relative to activations, a phenomenon which was referred to as "haemodynamic rectification". The notion that a single neuronal response may differ, in terms of the relative contributions of early and late neural components, was addressed finally. The results showed that the average BOLD response is attenuated in the context of large late or "endogenous" neuronal components. In contrast, as stipulated by the haemodynamic model, the rCBF response estimates were unaffected by changes in the experimental design parameters or the relative contribution of early and late neural components. The saturation effects observed in the BOLD response, but not in the rCBF response, were therefore attributed to the nonlinear transduction of rCBF to BOLD signal.

Chapter 5 focused on the impact of stimulus sequence (i.e. blocked vs. randomised) on the efficiency of response estimation in event-related fMRI. Blocked designs typically induce greater variance than randomised designs in the low frequency components that are "passed" by the HRF. This results in the standard error being smaller for contrasts testing for activations in blocked than randomised designs. As a result, significant effects were found in the blocked but not the randomised design in a number of regions and yet the amplitude of the haemodynamic responses for the two presentation modes was exactly the same. The results also showed that it may not be correct to assume that the error variance (i.e. the residual variance after evoked changes, modelled by the design matrix, have been discounted) is independent of changes in the experimental design. This makes *a priori* estimation of efficiency problematic, as the error variance must be taken into account but can only be estimated by performing a statistical analysis.

Chapter 6 investigated the neurophysiological effect of stimulus rate (i.e. the influence that stimulus rate exerts on neuronal responses) during reading and compared

the results from different scanning modalities. The effect of stimulus rate on neuronal activity was addressed by avoiding anecdotal comparison of the contrast-specific statistics and directly comparing the haemodynamic responses for different stimulus rates. This ensured that the observed differences were related to the underlying neuronal responses rather than the efficiency with which these responses were estimated. FMRI replicated the linear increases with stimulus rate seen in PET, but did not replicate the linear decreases. The failure to detect decreases in activity with fMRI is consistent with the haemodynamic rectification that was described in Chapter 4.

Chapter 7 turned to issues in data analysis and compared event-related and epoch models in the context of blocked design fMRI. It was demonstrated that, in a number of critical language regions, an event-related model explained changes in activity that were not accounted for by an epoch model. This means that the event-related analysis provided greater sensitivity to experimentally-induced effects than the epoch analysis. It was shown that the advantage of the event-related analysis was engendered by its early onset relative to the epoch model.

Finally, Chapters 8 and 9 presented two complementary analyses of the effects of word type (i.e. words vs. pseudowords) on brain activity, based on the principles of functional specialisation and functional integration. The results showed that pseudoword reading increases neuronal activity relative to word reading and fixation, and is associated with greater effective connectivity between regions than word reading. Furthermore, the two analyses identified a number of effects of word type at the individual subject level that were inconsistent across subjects. In this final chapter, I focus on three of the themes that have emerged from the experiments presented in this thesis. These include (i) the relationship between BOLD saturation, neurophysiological and efficiency-mediated effects, (ii) the differential sensitivity of event-related and epoch analyses in blocked design fMRI, and (iii) the inconsistencies in the effects of reading words vs. pseudowords across subjects. The chapter concludes with a section on issues that warrant further investigation.

10.2 BOLD saturation, efficiency-mediated and neurophysiological effects

In this thesis, I have shown that experimental design parameters may affect sensitivity in multiple ways as a consequence of BOLD saturation, efficiency-mediated and neurophysiological effects. Critically, these effects are not independent but interact as represented in Figure 10.1.



Figure 10.1 The relationship between BOLD saturation, neurophysiological and effeciencymediated effects.

The relationship between neurophysiological, BOLD saturation and efficiencymediated effects can be described as follows. First, neurophysiological effects may modulate the degree of saturation in the BOLD signal. This is illustrated in Chapter 4 of this thesis, in which increased neuronal activation is shown to reduce the BOLD response *per stimulus* or *event* even when the rCBF estimates remain constant (see Figure 4.5). Another example is given by haemodynamic rectification as discussed in Chapter 4. Here BOLD saturation effects are proportional to the degree of neuronal deactivation (see Figure 4.4).

Second, neurophysiological effects may contribute to the error variance (i.e. the residual variance after evoked changes, modelled by the design matrix, have been discounted), which in turn influences the efficiency with which parameters are estimated. For instance, neurophysiological effects that conform to the expected response, as modelled by the design matrix, will be associated with minimal error variance and parameters will be estimated efficiently. In contrast, neurophysiological effects that deviate from the expected response will induce greater error variance that may compromise design efficiency. An example is given in Chapter 5, in which greater error variance for randomised relative to blocked presentations was found in a number of frontal and occipital areas. It was suggested that differential error variance resulted from neurophysiological effects, i.e. differences in the shape and timing of the neuronal response due to different cognitive strategies for the two presentation modes.

Third, BOLD saturation effects may affect the efficiency with which effects are detected by contributing to the error variance. Specifically, the design efficiency will be reduced if strong BOLD nonlinearities occur that are not modelled by the design matrix.

In contrast, design efficiency will be preserved if nonlinearities in the BOLD signal are negligible or modelled appropriately.

A critical consequence of the observations above is that the multiple effects of design parameters should all be considered when constructing sensitive experimental designs. In practice, this may not be easy because, for example, the degree of BOLD saturation is likely to vary across different brain regions (Birn et al., 2001; Miezin et al., 2000). Furthermore, as demonstrated in Chapter 5, it may be problematic to estimate the relative efficiency of two or more experimental designs prior to statistical analysis. Disambiguating the source of stimulus rate effects after parameter estimation may also be difficult. For instance, while neurophysiological effects can be easily distinguished from efficiency-mediated effects by directly comparing different rates, it may not be possible to assess whether saturation effects observed at the haemodynamic level are due to BOLD nonlinearities or neurophysiological effects (e.g. neuronal habituation). This is because BOLD saturation and neurophysiological effects can only be segregated by acquiring the BOLD signal in conjunction with more direct measures of neuronal activity such as event-related potentials. Despite these difficulties, I hope that this thesis contributes to a better appreciation of how important it is to distinguish between BOLD saturation, efficiency-mediated and neurophysiological effects when (i) constructing sensitive designs and (ii) interpreting functional imaging results.

10.3 Comparing event-related and epoch models

Chapter 7 showed that an event-related statistical model may characterise the form (particularly the onset) of the observed haemodynamic response better than an epoch-

related model, even in the context of blocked design fMRI. Specifically, the eventrelated model appeared to be more sensitive to increases in activity for reading words and pseudowords relative to fixation. A question of interest is whether this finding can be generalised to (i) different cognitive paradigms and (ii) different experimental effects (e.g. reading 60 wpm vs. 20 wpm).

In fact it is possible that, when a different cognitive paradigm is used, an epoch model may provide greater sensitivity than an event-related model. For instance, neuronal activity may be sustained over time even if stimuli are presented intermittently. This would result in a steady-state haemodynamic response which is explained by an epoch model better than an event-related model. Even when neuronal activity is not sustained over time, an epoch analysis may still provide greater sensitivity than an event-related analysis. For instance, if the evoked haemodynamic response has an unusual (unpredicted) temporal shape, the epoch model is less likely to be out of phase and may be more accurate than an event-related model. An epoch analysis will also be more sensitive than an event-related analysis when the evoked haemodynamic response rises slowly or quickly returns to baseline. This is because epoch models reach their peak later and return to baseline earlier than event-related models, see Figures 1.3 and 7.1 in Chapters 1 and 7 respectively. In short, although Chapter 7 revealed greater sensitivity for the event-related than the epoch model, there are a number of cases when the evoked haemodynamic response may not conform to a series of delta functions. In these circumstances, an epoch model may be more accurate than an event-related model. Here a possible way of improving sensitivity in event-related analyses is to model

differences in the width and the onset time of the evoked haemodynamic response (e.g. using dispersion and time derivatives as implemented in SPM99 software).

The choice between an epoch and an event-related model may also depend on the research hypothesis that motivates the study. It is important to bear in mind that an epoch analysis models the mean activity within each block for each experimental condition. In contrast, the parameters of an event-related model encode the average response per stimulus or event for each experimental condition. This has practical implications when experimental conditions are compared that involve different stimulus rates, e.g. 20 wpm and 60 wpm. Here an epoch model contrast would compare the mean activity during blocks of 20 wpm to the mean activity during blocks of 60 wpm (as in PET). In contradistinction, an event-related contrast would compare the average response per stimulus for reading 20 wpm to the average response per stimulus for reading 60 wpm. As a result, the epoch model will indicate a significant effect of rate even when the response per stimulus or event is unaffected by rate whereas the eventrelated model will not. In contrast, the event-related model will indicate a significant effect of rate if the evoked haemodynamic response per stimulus or event is modulated by presentation rate. Figure 10.2 illustrates the counter-intuitive possibility that a positive effect of stimulus rate can be found using an epoch-related model, and yet a negative effect of stimulus rate may be detected when an event-related model is applied to the same data. As can be seen, the expected haemodynamic response to a train of stimuli is modulated by stimulus rate in the event-related but not the epoch model. Clearly, here the two analyses address different types of effects and the most appropriate model depends on the research question that motivates the study.



Figure 10.2 Simulated haemodynamic response (black line) to a train of stimuli when block length is 21 seconds and the SOA is either 3 seconds (left panels) or 1 second (right panels). Parameter estimates ($\hat{\beta}$) were obtained using an epoch model (red lines, upper panels) or an event-related model (blue lines, lower panels). The thin lines represent the underlying epoch and event-related neuronal models. In this simulation, the mean activity within the block increases as the number of stimuli increases or, equivalently, as the SOA decreases. As a result, when the epoch model was used, the parameter estimates increased with stimulus rate (i.e. from 0.5 to 0.8) (upper right). However, the increase in activity is not proportional to the number of stimuli being presented. Hence, the average evoked haemodynamic response *per stimulus* or *event* decreases as the number of stimuli increases. This is why, when the event-related model was used, the parameter estimates decreased with stimulus rate (i.e. from 1.7 to 0.9) (lower right).

To summarise, one of the contributions of this thesis was the demonstration that an event-related model may provide greater sensitivity than an epoch model. However, as discussed above, there may be a number of cases when the shape of the evoked haemodynamic response is such that an epoch model is more accurate than an event-related model. When experimental conditions are compared that involve different stimulus rates, the most appropriate model may depend on the research hypothesis.

10.4 Consistent and inconsistent effects across subjects

Chapters 8 and 9 addressed the impact of word type (i.e. words vs. pseudowords) on neuronal activity by looking at functional segregation and functional integration in the brain. Reading pseudowords relative to words increased neuronal responses in the left frontal operculum, left posterior inferior temporal gyrus and the right cerebellum (Chapter 8), and was associated with greater connectivity within a left-lateralized network including the frontal operculum, the inferior and superior temporal cortices, and the posterior fusiform gyrus (Chapter 9). Although these effects were relatively consistent across subjects, a number of effects were also found at the individual subject level that were not replicated across subjects. For instance, 7 out of 20 subjects showed significant but highly inconsistent increases in activity for words relative to pseudowords and fixation (see Table 8.4 in Chapter 8), and the impact of word type on the backward connection between the frontal operculum and the superior temporal cortex was significantly different across subjects (see Table 9.1 in Chapter 9).

As discussed in Chapters 8 and 9, there may be a number of reasons why functional imaging results are inconsistent across subjects. First, inter-subject variability

may result from different cognitive strategies adopted by the individual subjects. Second, inter-subject variability may occur when the experimental paradigm is relatively undemanding and leaves additional attentional resources that may be directed at other cognitive processes. Finally, differences in functional specialisation and integration may reflect the engagement of different brain regions even when the same cognitive strategy is adopted by all subjects.

Profound inter-subject variability may not only compromise sensitivity in random effect analyses (where the effect size is compared to the variability among subjects), but may also bias the results when the data are collapsed over subjects. For instance, significant results at the group level may be driven by strong effects in one or a subset of subjects only. Alternatively, null results at the group level may result from significant but highly variable (i.e. reverse) effects at the individual subject level. In this thesis, it has been proposed that inter-subject variability may account for some of the inconsistencies amongst previous functional imaging studies of word and pseudoword reading, which used fixed effect models and a limited number of subjects (see Chapter 8).

In short, the results presented in Chapters 8 and 9 suggest that inferences on the regional values averaged across subjects should be made with caution. Random effect analyses provide only a partial solution, as they allow consistent effects across subjects to be identified but are not sensitive to individual subject effects. The approach used in Chapter 8 involved the combination of random effect and individual subject analyses, which identified both consistent and inconsistent effects across subjects. Furthermore, in Chapter 9, an alternative analysis was presented which is one of the novel contributions

of the present thesis. This analysis involves the specification of a multi-subject network to investigate differences and commonalities across subjects in studies of functional integration.

10.5 Future work

There are a number of outstanding issues that require further investigation. Three of these will be discussed briefly in this section: (i) the coupling between neuronal activity and haemodynamics, (ii) the effect of the task on the neurophysiological effects of rate and duration, and (iii) the use of alternative analytical techniques for the estimation of effective connectivity.

10.5.1 The coupling between neuronal activity and haemodynamics

The biophysical simulations presented in Chapter 4 relied on an input-state-output haemodynamic model that was derived by combining the Balloon/Windkessel model (Buxton *et al.*, 1998) with a dynamical model of the transduction of neural activity into perfusion changes (Friston *et al.*, 2000). As discussed in Chapter 4, empirical characterization of the nonlinearities described by the study is required to further validate the findings and the model itself. This would require simultaneous measurement of neuronal activity and BOLD signal (i.e. see Logothetis *et al.*, 2001), or rCBF and BOLD signal (i.e. see Miller *et al.*, 2001), throughout the brain and across a range of sensorimotor and cognitive tasks. Preliminary results are consistent with a model that embodies a nonlinear step from stimulus to neuronal activity, a linear step

from neuronal activity to rCBF change, and a nonlinear step from CBF to BOLD signal change (Miller *et al.*, 2001).

10.5.2 The effect of the task on the neurophysiological effects of rate and duration

In Chapters 5 and 8, the neurophysiological effects of stimulus rate and duration were investigated using a single word and pseudoword reading paradigm. A number of regions were found throughout the brain that showed positive effects of rate and duration, that were consistent for words and pseudowords. A question of interest is whether these effects reflect intrinsic properties of specific brain regions and, as a result, are expressed across a wide range of language tasks. Alternatively, these effects may be context-specific and disappear when different language tasks are performed such as rhyming and lexical decision. To address this issue, a number of language tasks could be used that engage the same regions but rely on different cognitive processes. One possibility is that, in visual areas, the effects reported are expressed consistently over a range of tasks that involve visual stimulation. In contrast, some of the effects reported in temporal, parietal and frontal regions might be specific to single word reading.

10.5.3 Alternative analytical techniques for the estimation of effective connectivity

As discussed in Chapter 1, the evaluation of effective connectivity in the context of functional neuroimaging requires the use of a "causal model", in which regions and connections of interest are specified by the researcher. In this thesis, SEM was used to estimate the strength of the connections across different experimental conditions. Although SEM has been of great value to the neuroimaging community in the past few years, this approach is not without problems. For instance, SEM is used to make inferences about fMRI time series but it is not a time series model. This means that, if the time series were randomly permuted, SEM would give the same results. In other words SEM throws away temporal information, a deficiency that is not shared by other techniques such as Volterra kernels, Multivariate Autoregressor Modelling (MAR) and Dynamic Causal Modelling (DCM). Furthermore, SEM models that contain loops and self-connections generally do not provide unique estimators of the parameters. To overcome this problem, reciprocal connections were constrained to have the same values in the model used in Chapter 9. In contrast, alternative approaches such as Volterra kernels, MAR and DCM allow for unique estimation of the connection strength even in the presence of loops and self-connections. The reason for this is that these models do not contain instantaneous connections. Rather, effective connectivity arises from connections between regions at different time-lags. Further studies will therefore use techniques that exploit the temporal information embedded in the fMRI time series, and allow patterns of effective connectivity to be estimated even in the presence of loops and self-connections. A preliminary study comparing SEM, Volterra kernels and MAR suggests that the results obtained are consistent (Harrison et al., 2002).

A more fundamental issue relates to the biological parameters that should be used to make inferences about functional and effective connectivity. In previous PET and fMRI studies of functional integration, inferences have been made by looking at the coupling between haemodynamic responses in different brain regions. However, interactions in the brain do not occur at the level of haemodynamic responses, but at the neuronal level. Future work will employ models of functional and effective connectivity

that rely on the actual neuronal responses or well-constrained approximations to them. For instance, more direct measures of neuronal responses could be obtained using Electro-Encephalography (EEG) methods. Alternatively, a well-constrained approximation of the underlying neuronal activity could be obtained from the BOLD signal by deconvolution using an assumed haemodynamic response.

In conclusion, future studies will need to (i) empirically characterise the coupling between neuronal activity and perfusion changes in order to further validate the findings reported in this thesis and the haemodynamic model itself, (ii) investigate how context modulates the neurophysiological effects of stimulus rate and duration by employing a range of different language tasks, and (iii) use alternative analytical techniques to investigate the integration amongst functionally specialised regions in the brain.

Appendix

The General Linear Model and Efficiency

In the general linear model applied to fMRI time series, the response variable y is expressed in terms of a linear combination of explanatory variables in a design matrix X and a normally distributed error term N(0, $\sigma^2 \Sigma$). See Worsley and Friston (1995) for a treatment of the degrees of freedom that uses the Satterthwaite approximation to accommodate serial correlations when $\Sigma \neq I$.

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

$$\langle \mathbf{y} \rangle = \mathbf{X} \boldsymbol{\beta}$$
 (1a)

$$\operatorname{var}(\mathbf{y}) = \sigma^2 \mathbf{\Sigma} \tag{1b}$$

The parameter estimates $\hat{\beta}$ obtain in a least-squares sense using the pseudoinverse (denoted by +) of the design matrix

$$\hat{\boldsymbol{\beta}} = \mathbf{X}^{+} \mathbf{y} \tag{2}$$

$$\left\langle \hat{\boldsymbol{\beta}} \right\rangle = \mathbf{X}^{+} \left\langle \mathbf{y} \right\rangle = \boldsymbol{\beta}$$
 (2a)

$$\operatorname{var}(\hat{\boldsymbol{\beta}}) = \mathbf{X}^{+} \operatorname{var}(\mathbf{y}) \mathbf{X}^{+T} = \sigma^{2} \mathbf{X}^{+} \boldsymbol{\Sigma} \mathbf{X}^{+T}$$
(2b)

Inferences about effects of interest can be made using T statistics. An effect of interest is specified by a vector of contrast weights **c** that gives a weighted sum or compound of

parameter estimates $\mathbf{c}^{\mathrm{T}}\hat{\boldsymbol{\beta}}$, referred to as a contrast. The contrasts of parameter estimates $\mathbf{c}^{\mathrm{T}}\hat{\boldsymbol{\beta}}$ are an index of the size of the effects of interest. This is because the expected parameter estimates $\langle \hat{\boldsymbol{\beta}} \rangle$ are equal to the true parameters $\boldsymbol{\beta}$ (see 2a). The *T* statistic is simply the contrast divided by its estimated standard error, e.g. the square root of the contrast variance:

$$T = \mathbf{c}^{\mathrm{T}} \hat{\boldsymbol{\beta}} / (\operatorname{var}(\mathbf{c}^{\mathrm{T}} \hat{\boldsymbol{\beta}}))^{1/2}$$
(3)

where the contrast variance $var(\mathbf{c}^T \hat{\boldsymbol{\beta}})$ is a function of error variance σ^2 and the design matrix **X**. From (2b):

$$\operatorname{var}(\mathbf{c}^{\mathrm{T}}\hat{\boldsymbol{\beta}}) = \sigma^{2}\mathbf{c}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{X}^{\mathrm{+T}}\mathbf{c}$$
(4)

Here I refer to $\mathbf{c}^{T}\mathbf{X}^{T}\mathbf{\Sigma}\mathbf{X}^{+T}\mathbf{c}$ as design variance in contradistinction to $\sigma^{2}\mathbf{c}^{T}\mathbf{X}^{T}\mathbf{\Sigma}\mathbf{X}^{+T}\mathbf{c}$ which is contrast or estimator variance (note that when $\sigma^{2}=1$, the design and contrast variances are equivalent).

The contrast variance can be used to estimate the efficiency of an estimator for a specified contrast of interest c. Specifically, the efficiency of an estimator is inversely related to the contrast variance, e.g. it decreases with error variance σ^2 and design variance $c^T X^T \Sigma X^{+T} c$.

Appendix

$$(\text{Standard error})^2 = \frac{1}{\text{Efficiency}} = \sigma^2 \mathbf{c}^{\mathrm{T}} \mathbf{X}^{\mathrm{T}} \mathbf{\Sigma} \mathbf{X}^{\mathrm{+T}} \mathbf{c}$$
(5)

When contrast matrices are specified (e.g. when a number of basis functions are equally interesting), efficiency can be computed by using the trace operator (Dale, 1999):

$$\frac{1}{\text{Efficiency}} \propto \text{trace} \left[\mathbf{c}^{\mathrm{T}} \mathbf{X}^{\mathrm{T}} \mathbf{\Sigma} \mathbf{X}^{+\mathrm{T}} \mathbf{c} \right]$$
(6)

Clearly to compute the *T* statistic, standard error or efficiency one must estimate σ^2 . This is achieved through dividing the residual sum of squares by its expectation when $\sigma^2=1$

$$\langle \mathbf{r}^{\mathrm{T}} \mathbf{r} \rangle = \langle \operatorname{trace}(\mathbf{r} \mathbf{r}^{\mathrm{T}}) \rangle = \langle \operatorname{trace}(\mathbf{R} \mathrm{y} \mathrm{y}^{\mathrm{T}} \mathbf{R}^{\mathrm{T}}) \rangle = \sigma^{2} \operatorname{trace}(\mathbf{R} \Sigma \mathbf{R}^{\mathrm{T}}) = \sigma^{2} \operatorname{trace}(\mathbf{R} \Sigma)$$
(7)

$$\hat{\sigma}^2 = \frac{\mathbf{r}^{\mathrm{T}} \mathbf{r}}{\mathrm{trace} (\mathbf{R} \boldsymbol{\Sigma})}$$
(7a)

where \mathbf{r} is the vector of residuals and \mathbf{R} is the residual-forming matrix:

$$\mathbf{r} = \mathbf{R}\mathbf{y} \tag{7b}$$

$$\mathbf{R} = \mathbf{I} \cdot \mathbf{X} \mathbf{X}^{\mathrm{T}}$$
(7c)

References

References

Aertsen, A., and Preissl, D. H. (1991). Dynamics of activity and connectivity in physiological neuronal networks. In *Non Linear Dynamics and Neuronal Networks*. H.G. Schuster, New York pp. 281-302.

Aguirre, G. K., Zarahn, E., and D'Esposito, M. (1997). Empitical Analyses of BOLD fMRI Statistics. II. Spatially smoothed data collected under null-hypothesis and experimental conditions. *NeuroImage* 5: 199-212.

Ashburner, J., Neelin, P., Collins D. L., Evans, A., and Friston, K. J. (1997). Incorporating prior knowledge into image registration. *NeuroImage* 6: 344-352.

Bandettini, P. A., and Cox, W. (1998). Functional Contrast in Event-Related fMRI: Interstimulus Interval Dependency and Blocked Design Comparison. *NeuroImage* 7: S552.

Barker, A. T., Jalinous, R., and Freeston, I. L. (1985). Non-invasive magnetic stimulation of the human motor cortex. *Lancet* 1: 1106-1107.

Bentler, P. M. (1992). EQS Structural Equations Program Manual. Los Angeles: BMDP Statistical Software.

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Rao, S. M., and Cox, R.
W. (1999). Conceptual processing during the conscious resting state: a functional MRI study. *J. Cogn. Neurosci.* 112: 80-93.

Birn, R. M., Saad, Z., and Bandettini, P. A. (2001). Spatial heterogeneity of the nonlinear dynamics in the fMRI BOLD response. *NeuroImage* 14: 817-826.

Bloch, F., Hansen, W. W., and Packard, M. (1946). The nuclear induction experiment. *Phys. Rev.* **70**: 474-485.

Bokde, A. L., Tagamets, T.-A., Friedman, R. B., and Horwitz, B. (2001). Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron* **30**: 609-17.

Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.

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

Büchel, C., and Wise, R. J. S., Mummery, C. J., Poline, J.-B., and Friston, K. J. (1996). Nonlinear regression in parametric activation studies. *NeuroImage* **4**: 60-66.

Büchel, C., and Friston, K. J. (1997). Modulation of Connectivity in Visual Pathways by Attention: Cortical Interactions Evaluated with Structural Equation Modelling and fMRI. *Cerb. Cortex* 7: 768-778.

Büchel, C., and Friston, K. J. (1998). Dynamic Changes in Effective Connectivity Characterized by Variable Parameter Regression and Kalman Filtering. *Hum. Brain Mapp.* **6**: 403-408.

Büchel, C., Price, C. J., and Friston, K. J. (1998) A multimodal language region in the ventral visual pathway. *Nature* **394**: 274-277.

Büchel, C., and Friston, K. J. (2000). Assessing interactions among neuronal systems using functional neuroimaging. *Neural Networks* 13: 871-882.

Bullmore, E., Horwitz, B., Honey, G., Brammer, M., Williams, S., and Sharma, T. (2000). How Good is Good Enough in Path Analysis of fMRI Data? *NeuroImage* 11: 289-301.

Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., Koutstaal, W., Schacter, D., Rosen, B. R., and Dale, A. M. (1998). Functional-anatomic correlates of object priming in humans revealed by rapid presentation event-related fMRI. *Neuron* **20**: 285-296.

Buxton, R. B., Wong, E. C., and Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: The Balloon model. *Magn. Res. Med.* **39**: 855-864.

Cannestra, A. F., Pouratian, N., Shomer, M. H., and Toga, A. W. (1998). Refractory periods observed by intrinsic signal and fluorescent dye imaging. *J. Neurophysiol.* **80**: 1522-1532.

Cannestra, A. F., Bookheimer, S. Y., Pouratian, N., O'Farrell, A., Sicotte, N., Martin, N. A., Becker, D., Rubino, G., and Toga, A. W. (2000). Temporal and Topographical Characterization of Language Cortices Using Intraoperative Optical Intrinsic Signals. *NeuroImage* **12**: 41-54.

Chertkow, H., Bub, D., Waters, G., Evans, A., Whitehead, V., and Hosein, C. (1993). Separate effects of instructions and stimuli on cerebral blood flow: A ¹⁵O positron emission tomographic study. *Neurology* **43**: A189.

Clark, V. P., Maisog, J. M., and Haxby, J. V. (1998). fMRI study of face perception and memory using random stimulus sequences. *J. Neurophysiol.* **76**: 3257-3265.

Corbetta, M., Miezin, F. M., Dobmeyer, S., Shulman, F., and Petersen, S. E. (1991). J. Neurosci. 11: 1283-1302.

Coull, J. T., Büchel, K. J., Friston, K. J., and Frith, C. D. (1999). Noradrenergically Mediated Plasticity in a Human Attentional Neuronal Network. *NeuroImage* **10**: 705-715.

Dale, A. M., and Buckner, R. L. (1997). Selective averaging of rapidly presented individuals trials using fMRI. *Hum. Brain Mapp.* 5: 329-340.

Dale, A. M. 1999. Optimal experimental design for event-related fMRI. Hum. Brain Mapp. 8: 109-114.

Della-Maggiore, V., Sekuler, A. B., Grady, C. L., Bennett, P. J., Sekuler, R., and McIntosh, A. R. (2000). J. Neurosci. 20: 8410-8416.

Demonet, J. F., Price, C. J., Wise, R., and Frackowiak, R. S. J. (1994). Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: a positron emission tomography study. *Neuroscience Letters* **182**: 25-28.

Dettmers, C., Connelly, A., Stephan, K. M., Turner, R., Friston, K. J., Frackowiak, R. S. J., and Gadian, D. G. (1996). Quantitative Comparison of Functional Magnetic Resonance Imaging with Positron Emission Tomography using a Force-Related Paradigm. *NeuroImage* 4: 201-209.

Dhankhar, A., Wexler, B. E., Fulbright, R. K, Halwes, T., Blamire, A. M., and Shulman, R. G. (1997). Functional Magnetic Resonance Imaging Assessment of the Human Brain Auditory Cortex Response to Increasing Word Presentation Rates. *J. Neurophysiol.* 1: 476-483.

Duelli, R., and Kuschinsky, W. (1993) Changes in brain capillary diameter during hypocapnia and hypercapnia. J. Cereb. Blood Flow Metab. 13: 1025-1028.

Fiez, J. A., and Petersen, S. A. (1998). Neuroimaging studies of word reading. Proc. Natl. Acad. Sci. USA 95: 914-921.

Fiez, J. A., Balota, D. A., Raichle, M. E., and Petersen, S. E. (1999). Effects of Lexicality, Frequency, and Spelling-to-Sound Consistency on the Functional Anatomy of Reading. *Neuron* 24: 205-218.

Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J (1993a). Functional connectivity: The principal component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* **13**: 5-14.

Friston, K. J., Frith, C. D., and Frackowiak, R. S. J (1993b). Time-dependent changes in effective connectivity measured with PET. *Hum. Brain Mapp.* 1: 69-80.

Friston, K. J. (1995a). Regulation of rCBF by Diffusible Signals: An Analysis of Constrains on Diffusion and Elimination. *Hum. Brain Mapp.* **3**: 56-65.

Friston, K. J. (1995b). Functional and effective connectivity in neuroimaging: A synthesis. *Hum. Brain Mapp.* **2**: 56-78

Friston, K. J, Ashburner, J., Frith, C. D., Poline, J.-B., Heather, J. D., and Frackowiak,
R. S. J. (1995a). Spatial registration and normalisation of images. *Hum. Brain Mapp.* 2: 165-189.

Friston, K. J., Holmes, A. P., Worsely, K. J., Poline, J.-B., Frith, C. D., and Frackowiak,
R. S. J. (1995b). Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* 2: 189-210.

Friston, K. J., Büchel, C., Fink, G. R., Morris, J., Rolls, E., and Dolan R. J. (1997). Psychophysiological and Modulatory Interaction in Neuroimaging. *NeuroImage* 6: 218-229.

References

Friston, K. J., Josephs, O., Rees, G., and Turner, R. (1998). Nonlinear event-related responses in fMRI. *Magn. Res. Med.* **39**: 41-52.

Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N. A., and Dale, A. M. (1999a). Stochastic Designs in Event-Related fMRI. *NeuroImage* 10: 607-619.

Friston, K., J., Holmes, A. P., and Worsley, K. J. (1999b). How many subjects constitute a Study? *NeuroImage* **10**: 1-5.

Friston, K. J., Mechelli, A., Turner, R., and Price, C. J. (2000). Nonlinear responses in fMRI: The Balloon model, Volterra kernels and other hemodynamics. *NeuroImage* **12**: 466-477.

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

Friston, K. J., Penny, W., Phillips, C., Kiebel, S., Hinton, G., and Ashburner, J. (2002). Classical and Bayesian Inference in Neuroimaging: Applications. *NeuroImage* 16: 484-512

Frith, C. D., Friston, K. J., Liddle, P. F., and Frackowiak, R. S. J. (1991). A PET study of word finding. *Neuropsychologia* 29: 1137-1148.

Furey, M. L., Pietrini, P., Alexander, G. E., Schapiro, M. B., and Horwitz, B. (2000). Cholinergic enhancement improves performance on working memory by modulating the functional activity in dinstinc brain regions: a positron emission tomography regional blood flow study in healthy humans. *Brain Res. Bull.* **51**: 213-218.

Gally, J. A., Montague, P. R., Reeke, G. N., and Edelman, G. M. (1990). The NO hypothesis. *Proc. Natl. Acad. Sci. USA* 87: 6082-6086.

Gandhi, A. P., Heeger, D. J., and Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proc. Natl. Acad. Sci. USA* **96**: 3314-3319.

Gerstein, G. L., and Perkel, D. H. (1969). Simultaneously recorded trains of action potentials: analysis and functional interpretation. *Science* 164: 828-830.

Gonçalves, M. S., Hall, D. A., Johnsrude, I. S., and Haggard, M. P. (2001). Can meaningful effective connectivities be obtained between auditory cortical regions? *NeuroImage* 14: 1353-1360.

Gusnard, D. A., and Raichle, M. E. (2001). Searching for a baseline: functional imaging and the resting human brain. *Nature Reviews* **2**: 685-694.

Haase, A., Frahm, J., Matthaei, D., Hänicke, W., and Merboldt, K. D. (1986). FLASH imaging. Rapid NMR imaging using low flip-angle pulses. *J. Magn. Reson.* 67: 258-266.

Hagoort, P., Indefrey, P., Brown, C., Herzog, H., Steinmetz, H., and Seitz, R. (1999). The neural circuitry involved in the reading of german words and pseudowords: a PET study. *J. Cogn. Neurosci.* **11**: 383-398.

Harrison, L., Penny, W., and Friston, K. J. (2002) Multivariate Autoregressive Modelelling of fMRI time series. HBM abstract #416, Elsevier Science (USA). Available at <u>http://www.apnet.com/www/journal/hbm2002/14678.html#14678</u>

Hebb, D. O. (1949). The organization of behavior. New York: Wiley.

Heinze, H. J., Mangun, L. R., Burchert, W., Hinrichs, H., Scholz, M., Munte, T. F., Gos, A., Scherg, M., Johannes, S., and Hundeshagen, H. (1994). Combined spatial and temporal imaging of brain activity during visual selective attention in humans. *Nature* **372**: 543-546.
Henson, R. N. A., Price, C. J., Rugg, M. D., Turner, R., and Friston, K. J. (2002). Detecting Latency Differences in Event-Related BOLD Responses: Application to Words versus Nonwords and Initial versus Repeated Face Presentations. *NeuroImage* **15**: 83-97.

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

Higham, N. J. (1993). Optimization by direct search in matrix computations. SIAM J. Matrix Anal. Appl. 14: 317-333.

Horwitz, B., and Sporns, O. (1994). Neural modeling and functional neuroimaging. *Hum. Brain Mapp.* 1: 269-283.

Horwitz, B., Rumsey, J. M., and Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proc. Natl. Acad. Sci. USA* **95**: 8939-8944.

Horwitz, B., Tagamets, M.-A., and McIntosh, A. R. (1999). Neural modeling, functional brain imaging, and cognition. *Trends in Cognitive Sciences* **3**: 91-98.

Horwitz, B., Fiston, K. J., and Taylor, J. G. (2000). Neural modeling and functional brain imaging: an overview. *Neural Networks* 13: 829-846.

Huettel, S. A., and McCarthy, G. (2000). Evidence for a Refractory Period in the Hemodynamic Response to Visual Stimuli as Measured by MRI. *NeuroImage* **11**: 547-553.

Hutton, C., Howseman, A. M., Josephs, O., Zarahn, E., Friston, K. J. and Turner, R. (1998). The effect of Inter-Stimulus Interval on Signal Response in fMRI. *NeuroImage* 7: S591.

Hyder, F., Rothman, D. L., Mason, G. F., Rangarajan, A., Behar, K. L., and Shulman, R.
G. (1997). Oxidative glucose metabolism in rat brain during single forepaw stimulation:
a spatially localised 1H[13C] nuclear magnetic resonance imaging study. J. Cereb.
Blood Flow Metab. 17: 1040-1047.

Iadecola, C., Beitz, A. J., Renno, W., Xu, X., Mayer, B., and Zhang, F. (1993). Nitric oxide synthase-containing neural processes on large cerebral arteries and cerebral microvessels. *Brain Res.* **606**: 148-155.

Iidaka, T., Omori, O., Murata, T., Kosaka, H., Yonekura, Y., Okada, T., and Sadato, N. (2001). Neural Interaction of the Amygdala with the Prefrontal and Temporal Cortices in the Processing of Facial Expressions as Revealed by fMRI. *J. Cogn. Neurosci.* **13**: 1035-1047.

Jernigan, T. L., Ostergaard, A. L., Law, I., Svarer, C., Gerlach, C., and Paulson, O., (1998). Brain Activation during Word Identification and Word Recognition. *NeuroImage* 8: 93-105.

Jöreskog, K. G., and Sörbom, D. (1990). Model search with TETRAD II and LISREL. Sociological Methods and Research 19: 93-106.

Josephs, O., and Henson, R. N. A. (1999). Event-related fMRI: modelling, inference and optimisation. *Phil. Trans. R. Soc. Lond.* **354**: 1215-1228.

Josephs, O., Turner, R., and Friston, K. J. (1997). Event-related fMRI. Hum. Brain Mapp. 5: 243-248.

Kucera H., and Francis W. H. (1967). Computational analysis of present-day American English. Providence, Brown University Press.

Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* **297**: 203-205.

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weiskoff, R. M., Poncelet, B. P., Kennedy, D. N., Hoppel, B. E., Cohen, M. S., Turner, R., Cheng, H.-M., Brady, T. J., and Rosen, B. R. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. Sci. USA* **89**: 5675-5679.

Lashley, K. S. (1929). Brain mechanisms and intelligence: A quantitative study of injuries to the brain. New York: Dover Publications, Inc.

Lauterbur, P. C. (1973). Image formation by induced local interactions: examples employing nuclear magnetic resonance. *Nature* **242**: 190-191.

Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**: 150-154.

Maguire, E. A., Mummery, C. J., and Büchel, C. (2000). Patterns of Hippocampal-Cortical Interaction Dissociate Temporal Lobe Memory Subsystems. *Hippocampus* 10: 475-482.

Mandeville, J. B., Marota, J. J., Ayata, C., Zaharchuk, G., Moskovitz, M. A., Rosen, B. R., and Weisskoff, R. M. (1999). Evidence of a cerebrovascular postarteriole windkessel with delayed compliance. *J. Cereb. Blood Flow Metab.* **19**: 679-689.

Mansfield, P. (1977). Multi-planar image formation using NMR spin echoes. J. Physiol. 10: 349-352.

Marshall, J. C., and Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. *Journal of Psycholinguistic Research* **2**: 175-198.

Mazoyer, B., Zago, L., Mellet, E., Bricogne, S., Etard, O., Hounde, O., Crivello, F., Joliot, M., Petiti, L., and Tzourio-Mazoyer, N. (2001). Cortical networks for working memory and executive functions sustain the conscious resting state in man. *Brain Research Bulletin* 54: 287-98.

MacCallum, R. C. (1995). Model specification: Procedures, strategies and related issues. In *Structural Equation Modelling: Concepts, Issues and Applications*. R. H. Hoyle, Sage, Thousand Oaks pp.16-36.

McClelland, J. L., and Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. 1. An account of basic findings. *Psychological Reviews* 88: 375-407.

McCrory, E., Frith, U., Brunswick, N., and Price, C. J. (2000). Abnormal functional activation during a simple word repetition task: a PET study of adult dyslexics. *J. Cogn. Neurosci.* **12**: 752-761.

McIntosh, A. R., and Gonzalez-Lima, F. (1994). Structural equation modelling and its application to network analysis in functional brain imaging. *Hum. Brain Mapp.* **2**: 2-22.

McIntosh, A. R., Rajan, M. N., and Lobaugh, N. J. (1999). Interactions of prefrontal cortex related to awareness in sensory learning. *Science* **284**: 1531-1533.

McIntosh, A. R. (2000). Towards a network theory of cognition. *Neural Networks* 13: 861-70.

Mechelli, A., Friston, K. J., and Price, C. J. (2000). The effects of presentation rate during word and pseudoword reading: a comparison of PET and fMRI. *J. Cogn. Neurosci.* **12** Supplement 2:145-156.

Mechelli, A., Price, C. J., and Friston, K. J. (2001). Effective connectivity during reading: the effects of word type and stimulus rate. *NeuroImage* 13: S1292.

Menon, R. S., Ogawa, S., Kim, S.-G., Ellermann, J. M., Merkle, H., Tank, D. W. and Ugurbil, K. (1992). Functional Brain Mapping using MRI: Signal Changes Accompanying Visual Stimulation. *Invest. Radiol.* 27: S47-S53.

Miezin, F. M., Maccotta, L., Ollinger, J. M., Petersen, S. E., and Buckner, R. L. (2000). Characterizing the Hemodynamic Response: Effects of presentation Rate, Sampling Procedure, and Possibility of Ordering Brain Activity Based on Relative Timing. *NeuroImage* 11: 735-759.

Miller, K. L., Luh, W.-M., Liu, T. T., Martinez, A., Obata, T., Wong, E. C., Frank, L. R., and Buxton, R. B. (2001). Nonlinear Temporal Dynamics of the Cerebral Blood Flow Response. *Hum. Brain Mapp.* **13**: 1-12.

Mummery, C. J., Ashburner, J., Scott, S. K., and Wise, R. J. S. (1999). Functional neuroimaging of speech perception in six normal and two aphasic subjects. *J. Acoust. Soc. Am.* 106: 449-457.

Noppeney, U., and Price, C. J. (2002). A PET Study of Stimulus- and Task-Induced Semantic Processing. *NeuroImage* 15: 927-935.

Ogawa, S., and Lee, T.-M. (1990). Magnetic resonance imaging of blood vessels at high fields: *in vivo* and *in vitro* measurements and image simulation. *Magn. Res. Med.* 16: 9-18.

Ogawa, S., Lee, T.-M., Nayak, A. S., and Glynn, P. (1990). Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. *Magn. Res. Med.* 14: 68-78.

Ogawa, S., Tank, D. W., Menon, R., Ellerman, J. M., Kim, S.-G., Merkle, H., and Ugurbil, K. (1992). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc. Natl. Acad. Sci. USA* **89**: 5951-5955.

Orrison, L., Lewine, J., Sanders, J., and Hartshorne, M. (1995). Functional Brain Imaging. St Louis, Mosby.

Paradis, A.-L., Van de Morrtele, P.-F., Le Bihan, D., and Poline, J.-B. (1998). Do high temporal frequencies of the event-related fMRI response have a more specific spatial localization. *NeuroImage* 7: S606.

Paulesu, E., McCrory, E., Fazio, F., Menoncello, L., Brunswick, N., Cappa, S. F., Cotelli, M., Cossu, G., Corte, F., Lorusso, M., Pesenti, S., Gallagher, A., Perani, D., Price, C., Frith, C. D., and Frith, U. (2000). A cultural effect on brain function. *Nature Neuroscience* **3**: 91-96.

Pauling, L., and Coryall, C. D. (1936). The magnetic properties and structure of hemoglobin, oxyhemoglobin and carbonoxyhemoglobin. *Proc. Natl. Acad. Sci. USA* 22: 210-216.

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

Petersen, S. E., Fox, P. T., Synder, A. Z., and Raichle, M. E. (1990). Activation of extrastriate and frontal cortical areas by words and word-like stimuli. *Science* 249: 1041-1044.

Pollmann, S., Wiggins, C. J., Norris, D. J. von Cramon, D. Y., and Schubert, T. (1998). Use of Short Intertrial Intervals in Single-Trial Experiments: A 3T fMRI-Study. *NeuroImage* 8: 327-339.

221

Posner, M. I., Petersen, S. E., Fox, P. T., and Raichle, M. E. (1988). Localization of cognitive operations in the human brain *Science* 240: 1627-1631.

Price, C. J., Wise, R., Ramsay, S., Friston, K. J., Howard, D., Patterson, K., and Frackowiak, R. S. J. (1992). Regional response differences within the human auditory cortex when listening to words. *Neuroscience Letters* **146**: 179-182.

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

Price, C. J., Moore, C. J., and Frackowiak, R. S. J. (1996a). The effect of Varying Stimulus Rate and Duration on Brain Activity during reading. *NeuroImage* **3**: 40-52.

Price, C. J., Wise, R. J. S., and Frackowiak, R. S. J. (1996b). Demonstrating the Implicit Processing of Visually Presented Words and Pseudowords. *Cerebral Cortex* **6**: 62-70.

Price, C. J., and Friston, K. J. (1997). The temporal dynamics of reading: a PET study. *Proc. R. Soc. Lond. B* 264: 1785-1791.

Price, C. J., Veltman, D., Ashburner, J., Josephs, O., and Friston, K. (1999). The critical relationship between the timing of stimulus presentation and data acquisition in fMRI. *NeuroImage* **10**: 36-45.

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

Puce, A., Allison, T., and McCarthy, G. (1999). Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex* **9**: 445-458.

222

References

Purcell, E. M., Torrey, H. C., and Pound, R. V. (1945). Resonance absorption by nuclear magnetic moment. *Phys. Rev.* **69**: 37-38.

Rees, G., Howseman, A., Josephs, O., Frith, C. D., Friston, K. J., Frackowiak, R. S. J., and Turner, R. (1997). Characterizing the Relationship between BOLD Contrast and Regional Cerebral Blood Flow Measurements by Varying the Stimulus Presentation Rate. *NeuroImage* **6**: 270-278.

Ross, E. (1981). The aprosodias: Functional anatomic organisation of the affective components of the language in the right hemisphere. *Archives of Neurology* **38**: 561-564.

Rowe, J., Stephan, K. E., Friston, K. J., Frackowiak, R., Lees A., and Passingham, R. (2002). Attention to action in Parkinson's disease: impaired effective connectivity among frontal cortical regions. *Brain* **125**: 276-89.

Rugg, M. D., and Coles, M. G. H. (1995). *Electrophysiology of mind*. Chapter 1: pp. 1-17; Chapter 6: pp. 171-188.

Rumsey, J. M., Horwitz, B., Donohue, C., Nace, K., Maisog, J. M., and Andreason, P. (1997). Phonologic and orthographic components of word recognition: a PET-rCBF study. *Brain* **120**: 739-759.

Xu, B., Grafman, J., Gaillard, W. D., Ishii, K., Vega-Bermudez, F., Pietrini, P., Reeves-Tyler, P., DiCamillo, P., and Theodore, W. (2001). Conjoint and Extended Neural Networks for the Computation of Speech Codes: The neural Basis of Selective Impairment in Reading Words and Pseudowords. *Cerebral Cortex* **11**: 267-277.

Seidenberg, M. S., and McClelland, J. L. (1989). A distributed developmental model of word recognition and naming. *Psychological Reviews* **96**: 523-568.

Shallice, T. (1988). From neuropsychology to mental structure. New York: Cambridge University Press.

Sheskin, D. J. (1996). The handbook of parametric and Nonparametric statistical procedures. CRC Press.

Shulman, G. L., Corbetta, M., Buckner, R. L., Raichle, M. E., Fiez, J. A., Miezin, F. M., and Petersen, S. E. (1997). Top-Down Modulation of Early Sensory Cortex. *Cerebral Cortex* 7: 193-206.

Simos, P. G., Breier, J. I., Fletcher, J. M., Foorman, B. R., Castillo, E. M., and Papanicolaou, A. C. (2002). Brain Mechanisms for Reading Words and Pseudowords: an Integrated Approach. *Cerebral Cortex* **12**: 297-305.

Somers, D. C., Dale, A. M., Seiffert, A. E., and Tootell, R. B. H. (1999). Functional MRI reveals spatially specific attentional modulation in human primary cortex. *Proc. Natl. Acad. Sci. USA* **96**: 1663-1669.

St George, M., Kutas, M., Martinez, A., and Sereno, M. I. (1999). Semantic integration in reading: engagement of the right hemisphere during discourse processing. *Brain* 7: 1317-1325.

Toga, A. W., and Mazziotta, J. (1995). Brain Mapping The Methods. London: Academic Press.

Turner, R., Le Bihan, D., Moonen, C. T., Despres, D., and Frank, J. (1991). Echo-planar time course MRI of cat brain deoxygenation changes. *Magn. Res. Med.* **22**: 159-166.

Van Petten C., and Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Mem. Cognit.* 18: 380-93.

Vanzetta, I., and Grinvald, A. (1999) Increased cortical oxidative metabolism due to sensory stimulation: implications for functional brain imaging. *Science* **286**: 1555-1558.

Veltman, D. J., Friston, K. J., Sanders, G., and Price, C. J. (2000). Regionally specific sensitivity differences in fMRI and PET: Where do they come from? *NeuroImage* 11: 575-588.

Veltman, D. J., Mechelli, A., Friston, K. J., Price, C. J. (in press). The importance of distributed sampling in blocked fMRI designs. *NeuroImage*

Villringer, A., and Dirnagl, U. (1995). Coupling of brain activity and cerebral blood flow: basis of functional neuroimaging. *Cerebrovasc. Brain Metab. Rev.* 7: 240-276.

Walsh, V., and Rushworth, M. (1999). A primer of magnetic stimulation as a tool for neuropsychology. *Neuropsychologia* 37: 125-135.

Warrington, E. K., and McCarthy, R. A. (1987). Categories of knowledge: Further fractionation and an attempted integration. *Brain* **110**: 1273-1296.

Worsley K. J., and Friston, K. J. (1995). Analysis of fMRI time series revisited – again. *NeuroImage* 2: 173-181.

Vazquez, A. L., and Noll, D. C. (1998). Non-linear temporal aspects of the BOLD response in fMRI. *NeuroImage* 7: 108-118.

Zarahn, E., Aguirre, G. K., and D'Esposito, M. (1997). Empirical analyses of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *NeuroImage* 5: 179-197.