Functional Magnetic Resonance Imaging of Recovery from Post-Stroke Aphasia

Susannah Bethel, BSc.

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

March 2009

Abstract

This thesis presents the design, development and application of a novel overt picture-naming paradigm through a series of exploratory behavioural and imaging experiments. The paradigm is subsequently used in a functional magnetic resonance imaging study of recovery from post-stroke aphasia. The possibility of comparing correct and error naming responses in aphasic patients and unimpaired subjects induced to make errors was investigated and successfully trialled. This research improves on techniques currently favoured in imaging studies to explore the processes involved in functional recovery in a more analytical way. The novel study design provides a new way to interrogate processing involved in the production of aphasic responses.

The intentions of this project were to drive the research field of post-stroke aphasia recovery forward by suggesting and applying new methods of using functional imaging to investigate the current pertinent research questions. In addition to this, it was aimed that data collected from participants who have an aphasic deficit, and those with a healthy language system, would be analysed to provide evidence of how a stroke damaged brain may recover functional language. It was hypothesised that results from aphasic patients would show that successful language performance is associated with cortical activation of the patients' normal left hemispheric language areas, around their lesion site. Conversely, the hypotheses state that production of linguistic errors would correlate with an increase in activation in areas of the right hemisphere homologous to the left lateralised fronto-temporal language production network. It was thought that further investigation of successful and unsuccessful language performance in unimpaired speakers would echo this finding.

The current debate in this research field centres on the role of the undamaged hemisphere in successful recovery. Five chronic stage aphasics were tested using the developed continuous scanning, event-related paradigm and their correct and error naming trials were compared. Results indicate that recruitment of cortical areas homologous to the stroke lesion can support successful language processing. This is contrary to the theory that disinhibition of non-dominant language areas may contribute to the production of aphasic errors.

An investigation of forced errors in unimpaired speakers was also conducted to provide comparisons with the aphasic patient group. Imaging results showed that the naming-to-deadline paradigm used may provide a useful baseline for the normal processes involved in the monitoring and control of task performance.

Acknowledgements

I would like to thank the Stroke Association for supporting me personally and the research project. Thank you to Professor Philip Bath who housed me at the Division of Stroke Medicine at the beginning and to my supervisor, Dr Alan Sunderland, for his patience and lovely Sunday dinners. I would like to express my appreciation to Kay Head, Dr Susan Francis and Dr Penny Gowland for their technical advice on all fMRI issues.

I am especially grateful to my family for their support and encouragement, especially my long-suffering husband Carl, whose magic with computers always saves the day, and our cute babies, Bethan and Harri. As much as I love them, a big thank you has to go to the people who helped me with childcare when I needed it the most; my parents, my parents-in-law, Carl and Sharon across the road.

1 Introduction

1.1 Stroke Prevalence and Aetiology

The brain is especially vulnerable to injury resulting from blood flow abnormalities. A human brain makes up just 2% of overall body mass and yet up to 20% of cardiac output is directed there to meet its continuous need for glucose and oxygen (Siesjo, 1984). Any interruption to this supply, as happens in stroke, can result in localised or more widespread tissue death.

A stroke can be thought of as a brain attack and has two major causes. An ischaemic stroke is the most common type and is caused by a blockage of a blood vessel in the brain. This can happen through the build up of a clot inside the blood vessel (thrombosis), or the movement of a blockage from a different part of the body via the bloodstream to the brain (embolism). A haemorrhagic stroke occurs when a blood vessel either within the brain (intracerebral haemorrhage) or on the brain's surface (subarachnoid haemorrhage) bursts and blood is lost into the surrounding tissue.

Stroke is the single most common cause of adult disability in the western world (Feigin et al, 2003), with a third of all stroke patients suffering from persistent physical or cognitive deficits. It can occur at any age, although the risk of stroke rises dramatically as a person grows older. At the age of 85, one in four men and one in five women can expect to have a stroke (Wolfe et al). The Stroke Association reports that within the United Kingdom an estimated 150,000 people suffer a stroke every year. This figure equates to one stroke every five minutes (The Stroke Association, 2007). With people now living longer than ever before the incidence rates of stroke are likely to rise. With this in mind, researchers and rehabilitation specialists are working to determine what mechanisms the brain naturally employs in recovery from stroke and how they can be enhanced using therapy, to promote functional recovery in patients.

1.2 Aphasia in Stroke

There is a range of potential physical and cognitive deficits facing a patient recovering from stroke. A particularly potent problem amongst this patient population is that of aphasia. Aphasia can be described as an umbrella term for deficits in language production and comprehension as a result of brain damage. The specific linguistic problems faced by each aphasic individual can vary widely and presents a challenge to researchers and therapists searching for successful rehabilitative techniques for all patients. Aphasia is, however, a significant area of stroke research due to its impact on this patient population. Studies have utilised a variety of experimental designs to research the prevalence of aphasia in stroke patients and have reported incidence rates ranging from 21% to 38% (Engelter et al, 2006).

Aside from the obvious direct impact a communication impairment has on a person's day-to-day living and relationships, aphasia has also been associated with an increased presence of depressive symptoms in stroke patients. Kauhanen et al (2000) report that at 3 months post-stroke 70% of aphasic patients presented with depression, falling slightly to 62% at 12 months. The proportion of these patients who were diagnosed with major depression, however, increased over the 12 month post-stroke period from 11% to 33%. Similarly, Medina et al (2007) found that patients with primary progressive aphasia, where language declines steadily over time, report more depressive symptoms than controls and, specifically, that the number of reported symptoms correlated with the degree of naming impairment suffered by patients. In their longitudinal study of post-stroke depression Thomas and Lincoln (2006) studied the factors that predicted the severity of stroke patients' depression. Their results showed that greater communication impairment was predictive of severe depression in patients at recruitment into the study and associated with an increased likelihood of depression still being present at follow-up. In this study communication impairment was found to be the strongest predictor of the severity of a stroke patient's depression and its prognosis.

Aphasic stroke patients are also less likely to return to work than their non-aphasic counterparts (Blackschaffer & Osberg, 1990) and there is evidence that caregivers of aphasic stroke survivors report more difficulty with tasks and more negative stroke-related outcomes compared to caregivers of non-aphasic stroke patients (Bakas et al, 2006). All of the research mentioned above point to the burden aphasia places on the lives of sufferers

and their carers and underlines the need for this condition to be understood and effectively treated to enhance a patient's recovery from stroke and their ability to regain their former quality of life and relationships.

1.3 Neurology of Language and Aphasia

Language has been found to be lateralised to one dominant hemisphere in the majority of cases. Rasmussen and Milner (1977) used the Wada test, also known as the intracarotid amobarbital procedure, in a study to provide evidence of speech laterality in healthy volunteers. The Wada test works by injecting the person with a drug that temporarily deactivates the hemisphere on the side the injection has been administered. Subjects are asked to count backwards continuously. If their counting is disrupted for a period from approximately one to three minutes then the experimenter will conclude that the language dominant hemisphere is the side that was injected. In their study Rasmussen and Milner found that 96 percent of right-handed people are left dominant for language leaving just 4 percent with their speech lateralised in the right hemisphere. Left-handed people and those with mixed handedness were found to be left dominant 70 percent of the time. Out of the remaining 30 percent, 15 percent had right hemisphere lateralisation and 15 percent had their language organised bilaterally. Results from studies using neuroimaging technologies agree with these figures, reporting approximately 95% of right-handers showing left hemisphere language dominance with 20-27% of left-handers having right-lateralised language (Szaflarski et al, 2006).

As can be expected from evidence such as this, in the majority of cases aphasia has been found to be a result of damage to the left hemisphere of the brain.

1.3.1 The classical view

During the late nineteenth century autopsy work with patients who had suffered from speech and language impairment led to the discovery of Broca's area (1861) and Wernicke's area (1874) in the brain (see figure 1). These areas were sited in the inferior frontal gyrus and the superior temporal gyrus respectively, connected to each other by the arcuate fasciculus, and were identified as being central to the functioning of expressive and receptive speech. Early localisation theories explaining the causes of aphasia focused on damage to these two brain areas or the connections between them (Geschwind, 1970).

Using this simple model of the breakdown of language, lesions to specific parts of this network were thought to cause particular types of aphasia with their associated symptoms. For example, Broca's aphasia, a disorder of speech production, was so named because of the observation that it often coincided with damage to Broca's area. Similarly, damage to Wernicke's area was thought to predict the comprehension difficulties and fluent production of jargon associated with Wernicke's aphasia.

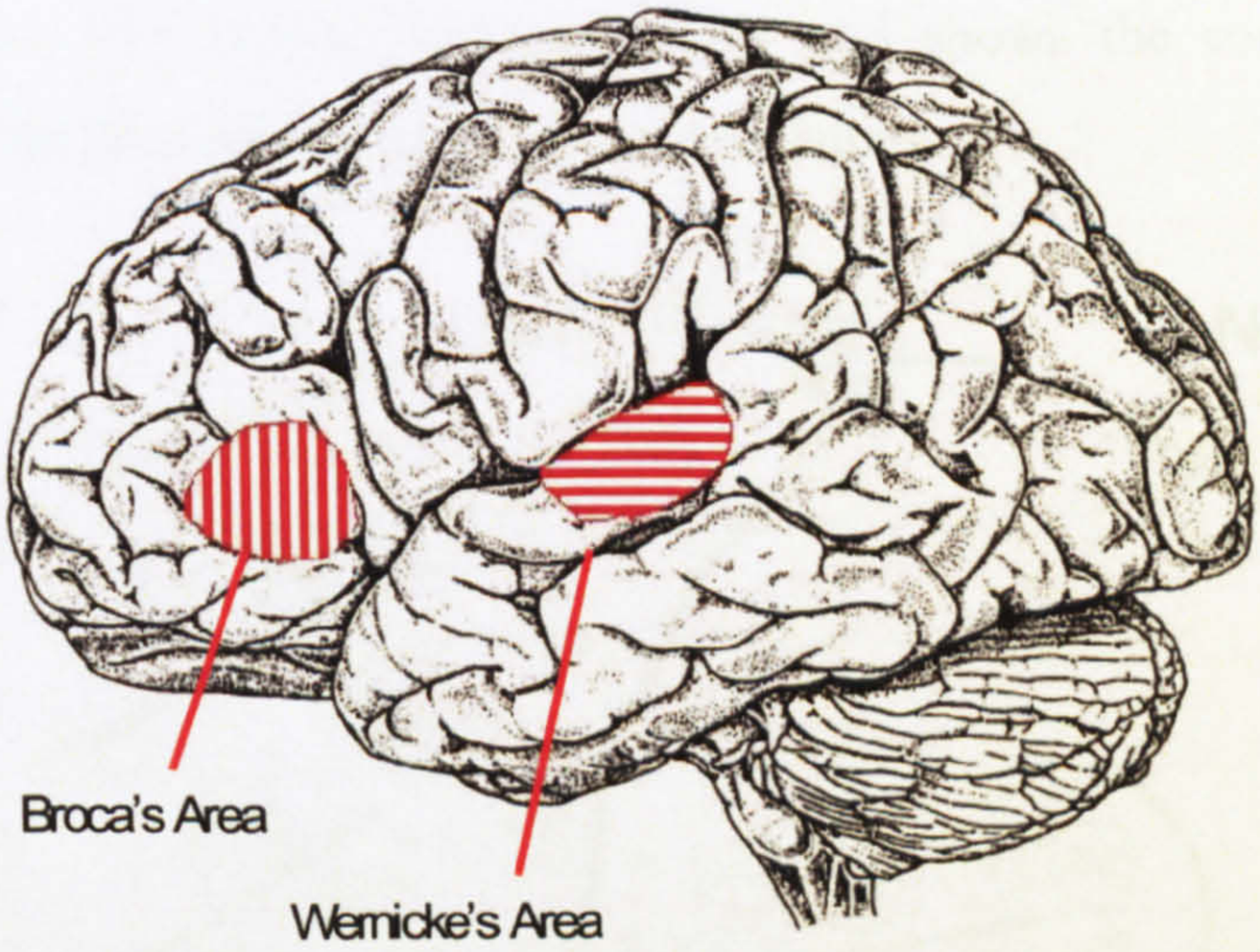

Figure 1. Approximate locations of Broca's area (Brodmann Area 44/45) and Wernicke's area (Broadmann Area 22) in the brain.

Although these terms are all still in use today the contemporary view of language organisation is much more distributed and it is recognised that a patient's aphasic symptoms rarely fit into these predefined classifications of aphasia types. For instance, Willmes and Poeck (1993) found in a retrospective study of 221 aphasic patients that only 35% of those with damage to Broca's area had the type of aphasia classically thought to be analogous to this region. Similarly, just 59% of the patient group with Broca's aphasia presented with damage to Broca's area. Dronkers et al (2000) corroborated these results in their studies, finding that only 50-60% of their patients with Broca's area lesions display persistent Broca's aphasia and conversely 15% of patients with Broca's aphasia have their Broca's area intact.

Despite these inconsistencies the classical model should, however, be credited with guiding research and clinical practice and attempting to provide an overall functional and anatomical model of language.

1.3.2 The contemporary view

Language is seen as a distributed network of components anatomically arranged across a bilateral network with left dominance (see Chapter 6 for a discussion of contemporary psycholinguistic models). A study by Ojemann et al (1989) serves as a neat demonstration of the variable cortical areas that are now implicated in language processing. The diagram below (figure 2) was compiled from an intraoperative electrical stimulation study of 117 epileptic patients with normal language ability and shows the cortical sites that when stimulated consistently induced errors in object naming.

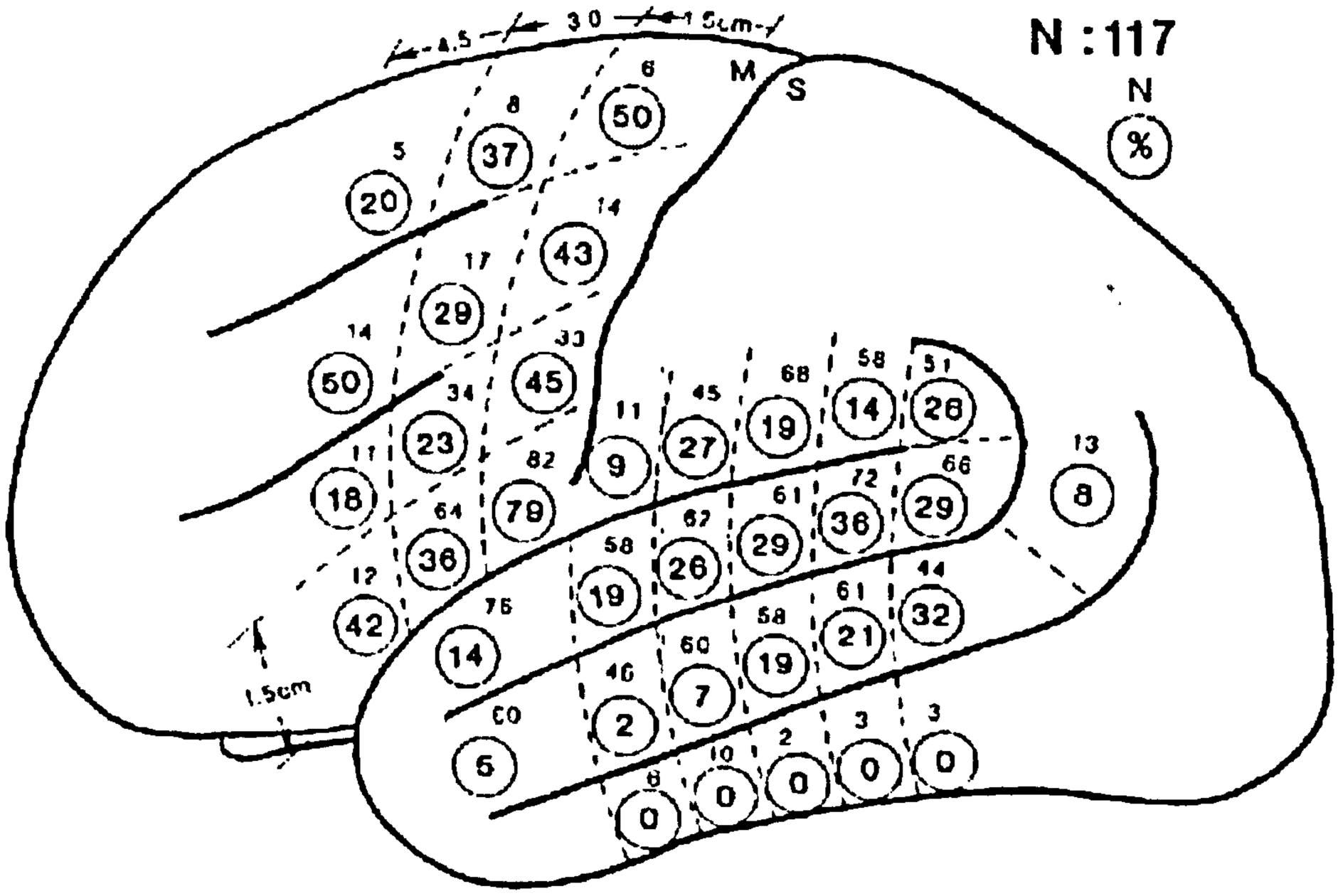

Figure 2. Diagram from Ojemann et al (1989) depicting the cortical sites deemed important for object naming. The brain has been divided into zones. The numbers above the circles in each zone are how many subjects had a language site in this zone. The circled numbers are the percentage of those subjects in whom stimulation provoked naming errors.

Recent neuroimaging studies also highlight the distributed nature of language processing (e.g. Indefrey & Levelt, 2004; Wise, 2003; Hillis, 2007).

Research now tends to focus more on restricted aspects of language function, trying to explain tiny pieces of the puzzle at any one time rather than looking at the broad view (Poeppel & Hickok, 2004). With the advances in psycholinguistics over the last century

language is known to be much more complex than was acknowledged in Broca and Wernicke's time. This necessarily has to inform and constrain neuropsychological theories and experimentation, although a gulf remains between the two sciences. It is often difficult to relate psycholinguistic theories to evidence of what occurs in the brain during normal and abnormal language processing and vice versa.

In recent times the development of new technologies and techniques has allowed the study of the living, working brain in both healthy and impaired people. The explosion of research using non-invasive, or minimally invasive, neuroimaging methods has especially contributed to the functional mapping of language and its breakdown in aphasia.

1.4 Functional Imaging in Stroke Research

Functional imaging techniques have opened up a new world for research into functions and dysfunctions of the brain. Unlike the traditional lesion-deficit studies, with which the discipline of neuropsychology established itself, imaging studies allow the brain to be viewed as a whole and to analyse the regions and functions of interest in the context of their positions and connections with other neural tissue.

Functional magnetic resonance imaging (fMRI) is a totally non-invasive method for viewing the brain in a working state. By imaging the brain whilst a subject is conducting a particular task, researchers can infer the specific neural areas that are involved in the completion of this task. In this way functional maps of the brain can be produced, based on the measures taken from the physiological responses of the brain to activation.

Experiments conducted with psychologically impaired patients can reveal much about the changes that occur in the brain after it is damaged. Brain areas that are necessary for a function can be determined and layers of redundancy can be teased out. For instance, if patients are seen to be performing on a task just as healthy control subjects would, but are not activating a particular region that control subjects do, then it can be concluded that this region is not necessary for the completion of this task (Price & Friston, 1999). Any compensatory changes in activation levels in other, perhaps remote, parts of the brain can also be identified and investigated. Figure 3 shows the possible explanations for abnormal activations in patients.

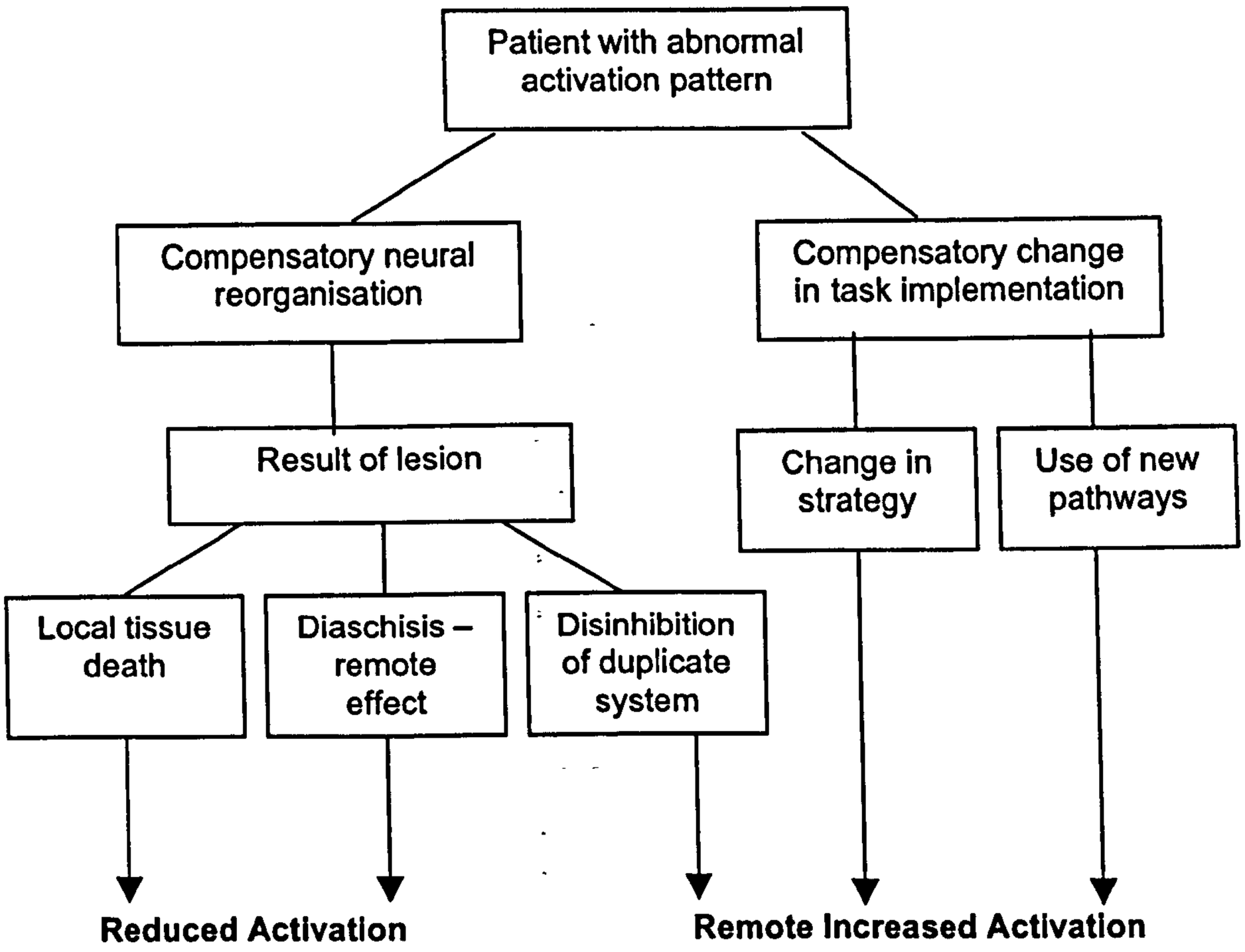

Figure 3. Damage induced changes in the brain that can be investigated with functional imaging (based on Price and Friston, 1999).

Neuroimaging is a powerful tool to use in stroke research, however, to maximise the benefit of knowledge gained from imaging studies researchers should combine them with other experimental methods. Evidence from functional imaging can only answer certain experimental questions and there is a danger that misleading conclusions can be drawn from these studies. Price and Friston (1999) highlighted this risk in their review of the types of imaging studies that can be conducted with patients. They pointed out that in order to properly investigate the changing neural architecture after the brain is lesioned, researchers need to make the distinction between neural reorganisation and cognitive reorganisation. To make valid inferences about changes in neural organisation we need to rule out the possibility of the cause being alterations in cognitive strategy. An abnormal activation pattern, compared to that of healthy controls, can arise due to the patient performing the task in a different way, perhaps to compensate for processes no longer available to them. Functional imaging data alone cannot determine if this is the case or not. Researchers need to be aware of this possible confound when interpreting results and,

if possible, should integrate their findings with those from behavioural testing of their patients to tease out the truth behind the evidence.

1.5 Mechanisms for Recovery

Many patients show good return of function after damage, however it is not yet clear exactly what mechanisms support this recovery. Researchers have reported a number of observations when studying the recovering brain that may be plausible candidates. The obvious decrease of functional activation within the lesioned areas has been linked to a significant increase in activation across other areas of the brain, usually within the contralateral hemisphere. This phenomenon has been demonstrated by most experiments conducted in this area. For example, Netz et al (1997) compared the motor output of stroke patients who have hemiparesis, a paralysis affecting one side of their body, with healthy control subjects. They elicited motor responses in all subjects using transcranial magnetic stimulation (TMS) and noted that the stroke patients showed different patterns of responses than the control group, which included a certain amount of reorganisation within the unaffected hemisphere.

Differences in activation levels within brain structures far removed from the functional network affected by damage have also been recorded. Seitz et al (1999) studied stroke patients who had recovered from a severe hemiparesis during performance of a finger movement task using their recovered hand. Measures of regional cerebral blood flow (rCBF) were obtained using a PET scanner. Subsequent comparisons of activation patterns between the patient group and a control group revealed metabolic change in areas that were remote to the lesion site, but affected by the ischemic injury. This occurrence has been given the term "diaschisis" and appears due to connections, anatomically and functionally, that exist between the lesion site and the distant areas.

A third change observed in a lesioned brain is the reactivation of perilesional tissue. This has often been seen to occur later in the recovery process, compared to reorganisation of activation and the appearance of diaschisis (Saur et al, 2006).

Research into recovery from aphasia has focused on two of the above phenomena. Specifically, the major debate in this area has concentrated on the role of the non-language

dominant hemisphere in recovery and what effect disinhibition of the right hemisphere after stroke damage has on aphasic language performance.

1.5.1 Disinhibition of right hemisphere

There are now a great number of studies reporting an increase of activation in the right hemisphere in stroke patients who have aphasia (Abo et al, 2004; Rosen et al, 2000; Sharp et al, 2004). This effect is particularly potent in the right hemisphere homologues of the classical language areas (Blank et al, 2003). For example, Gold et al (2000) compared a patient who had suffered extensive damage to the posterior frontal, temporal and parietal cortices in their left hemisphere with a healthy control. Both subjects were imaged whilst successfully performing the same language tasks. The control subject, as predicted by previous studies into normal language functioning, displayed a left-lateralised pattern of activation. This included the inferior frontal gyrus, the inferior and middle temporal gyri and the left angular gyrus. In contrast, the patient showed activation in areas such as the superior temporal gyrus, the middle frontal gyrus, the supramarginal gyrus and the angular gyrus, all in the right hemisphere. Basso et al (1989) produced case studies of patients with left hemisphere lesions who had made a partial recovery from aphasia only to relapse following subsequent right hemispheric lesions, indicating that for these patients the right hemisphere had been supporting their language recovery.

There is also evidence that the brain can be actively encouraged to utilise contralateral areas to provide an improvement in language task performance. Musso et al (1999) set out to achieve this in their research into the effects of training on patients with aphasia. Using PET they investigated what short-term changes could be made to the neural network involved in language comprehension in four patients diagnosed with comprehension difficulties. In this study, twelve consecutive measures of regional cerebral blood flow were taken whilst the patients were engaged in the activation task. This task consisted of understanding the difference between two different verbal commands and making appropriate responses. Between each of the twelve scans there was an interval of twelve minutes, in which patients underwent a comprehension test and an eight minute intensive training session. All four patients showed improvement in performance over the course of the experiment. The scores of the comprehension tests, administered between each scan and training session, were correlated with the brain areas that received increased blood flow

during the activation task. This revealed a modification of the activation pattern due to the elicited short-term recovery of language comprehension, indicated by an increase in the superior temporal, the middle temporal and the supramarginal gyri in the right hemisphere. This study serves to add more evidence to the theory of recovery from aphasia resulting from a shift of activation to homologous right hemispheric areas that are capable of compensating for the functional loss of the language dominant regions of the left. It also implicates an important role for rehabilitation in recovery, by demonstrating how the brain can be coerced into using this right-sided recovery network.

It would be tempting to conclude on the strength of this evidence that the recruitment of homologous areas of the contralateral hemisphere is a credible and effective mechanism for recovery from aphasia. However, as neat as this theory seems, it does appear to be a rather simplistic conclusion of what occurs in recovery from aphasia.

1.5.2 Re-emergence of perilesional activation

Recent studies have shown that an increase in activation within the right hemisphere may not be such a positive indicator of functional recovery at all. Rosen et al (2000) presented the now common finding of a stronger than normal response in the right hemisphere when imaging aphasic patients in varying states of recovery. However, they discovered that the level of activation seen in right hemispheric areas was not correlated to the level of performance displayed by the patients, whilst conducting a battery of lexical tasks. In addition to this, the researchers noticed some residual activation around the lesion site within the left hemisphere in two of the patients imaged. These two patients gave the best task performance of the subject group and also were seen to recover more completely from their aphasia.

The activation seen in the right hemisphere was interpreted by Rosen et al to reflect the loss of inhibition from the homologous areas in the left, or the use of an inefficient task strategy that involved the weaker areas of the non-dominant hemisphere. They believe that their study gives an insight into the functional importance to recovery of the reactivation of perilesional tissue and indicates the role of neural reorganisation in recovery to be more interfering than compensatory.

This conclusion has been supported by other published work (e.g. Warburton et al, 1999). Notably Karbe et al (1998), who showed clearly from correlations between task-related metabolic increases and number of errors made on the Token test, that the more persisting activation a patient had in right hemispheric areas then the more errors they tended to produce. They also demonstrated the converse to be true for a task-related increase in structurally recovered left hemispheric areas (see figure 4).

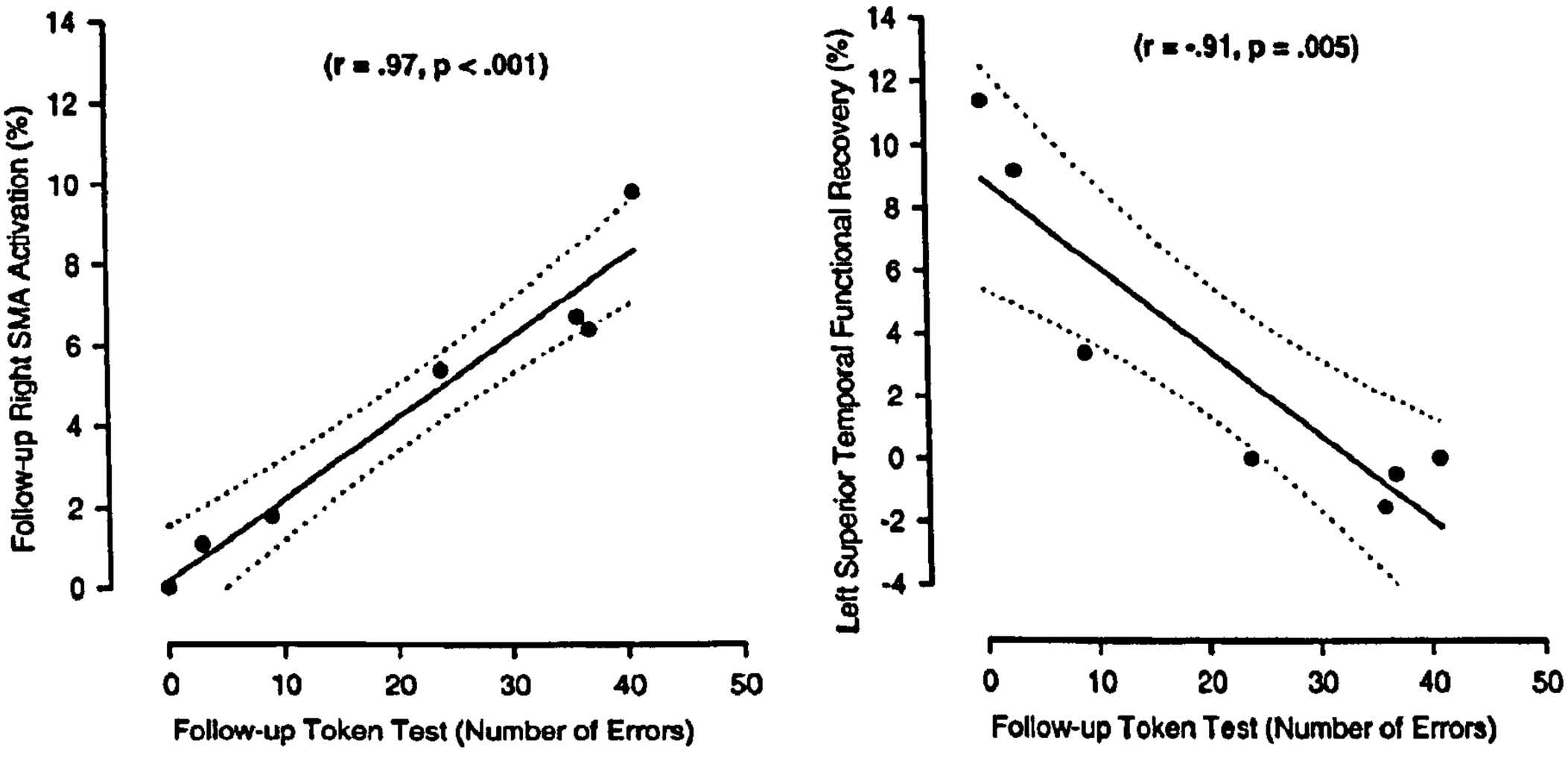

Figure 4. Correlation plots of increases in regional cerebral blood flow with number of errors on the Token test (from Karbe et al, 1998).

In their treatment study, Cornelissen et al (2003) imaged three anomic patients whilst conducting a picture naming task before and after training that focussed on their word-finding difficulties. All patients showed improvement in naming performance and the authors noted that the only area to show an associated increase in activity was located close to the lesion sites in the left inferior parietal lobe. This study provides more evidence of the important of reactivating left hemisphere perilesional areas in rehabilitation.

Transcranial Magnetic Stimulation (TMS) has been put to good use in this research area as a modern and more focused Wada test. Naeser et al (2005) used this technology to repetitively suppress a portion of the right Broca's area over a period of ten days in a group of chronic aphasics. Assessed picture naming ability showed post-TMS improvement that was still seen 8 months later in the majority of patients. This finding could be interpreted as a beneficial long-term re-inhibition of the right Broca's homologue, or a change from a

maladaptive naming strategy to a more effective one. A combination of TMS treatment and fMRI would be useful to determine what neural areas were supporting this improvement. It should be noted, however, that this study applied repetitive TMS specifically to the anterior portion of the right Broca's area, known as the right pars triangularis. In a previous TMS study Naeser et al (reviewed in Naeser et al, 2005) observed the opposite effect when suppressing activity in the posterior part of right Broca's area, the right pars opercularis. In this case the number of successfully named pictures by chronic nonfluent aphasic patients significantly decreased, coupled with an increase in naming latencies (Crosson et al, 2007). This suggests that a more complex view of recovery from aphasia is needed, rather than a straightforward left versus right battle.

1.5.3 A tale of two sides

Some studies have suggested that the specific areas that are important for recovery of language after stroke can vary depending on the site and size of the lesion and the time that has elapsed since the insult occurred. Work by Heiss et al (1997, 1999) and Vitali et al (2007) has shown that reorganisation of language to right hemisphere areas is associated with larger stroke lesions and a more severe aphasic deficit. It was found that better language recovery occurs when lesions are small and the normal left hemisphere areas are spared and reactivate to support language function. While this conclusion may seem intuitive it does suggest a hierarchy of areas that can contribute more or less effectively to the recovery of language function.

Longitudinal studies of aphasic patients have also implicated a role for both hemispheres in recovery from aphasia. The general finding has been that in the acute stages of recovery the right hemisphere increases its contribution to language processing with the left hemisphere showing a resurgence of activity in the chronic stage (Ansaldo and Arguin, 2003; Saur et al, 2006). Saur et al suggest that this pattern represents three distinct phases in recovery from aphasia; an immediate reduction in left lateralised activation followed by a release of right hemisphere language homologues that are capable of promoting language functional improvement. Finally, a shift back to left hemispheric dominance occurs to support further language recovery.

Taking all of the evidence presented above into consideration it cannot be denied that the non-dominant right hemisphere does have a part to play in the recovery process of aphasic patients. However, the pertinent question for researchers to address is what is the exact nature of the role of the right hemisphere? Does the often observed increase in activation represent a functionally significant recruitment of language homologues that are able to support recovery, or are researchers merely seeing the release of non-dominant areas through transcallosal disinhibition? It may be that these areas interfere with successful processing, or can contribute to language function but with reduced effectiveness than that of the lesioned areas. It is also possible that for some patients, with a certain set of circumstances, a shift of activation to the contralateral hemisphere is essential for any recovery from post-stroke aphasia to take place. The evidence, as it exists in the current literature, can support all of these various conclusions.

For the development of new rehabilitative or medical interventions it is crucial that the true mechanisms that support and enhance functional recovery in aphasic patients are determined. The future of research into aphasia recovery lies in the development of new, analytical experimental methods that can penetrate to the heart of the questions raised by the prevailing body of study results.

1.6 Recent Developments in the Literature

Most fMRI studies of language and aphasia have requested covert task responses from their participants, so that responses are merely thought of rather than vocalised (e.g. Ellis et al, 2006; Perani et al, 2003). This technique has been widely adopted due to the technical difficulties associated with acquiring usable data whilst speech is occurring (see Chapter 2 for a detailed discussion of the issues). Some studies are now beginning to be reported that record overt responses from healthy participants, extending the range of research questions that can now be interrogated with fMRI. Sparse or clustered imaging methods have often been used in overt studies to allow for the clear recording of responses without the interference of scanner noise in between image acquisitions (e.g. Abrahams et al, 2003; Barch et al, 2000). Use of these methods though is difficult when studying aphasic language production due to the relative unpredictability of the timing of responses.

A natural progression from the implementation of overt language paradigms in fMRI studies is to use the content of the responses as a basis for the data analysis. Aphasic participants may make errors or omit responses altogether. Now successful methods for imaging spoken responses are increasingly being used it does not seem a particularly valid research strategy to lump all responses together regardless of their quality. Different brain systems are likely to be involved in the production of correct and incorrect responses. If this is true then analysing all responses together may decrease the sensitivity to detect task-related haemodynamic changes. Also, now that there exists a wealth of evidence from studies simply comparing aphasic and normal language it seems there is a real need to investigate further using more penetrative methods, such as allowing for the contrast of an aphasic patient's task response types. In this way participants' responses can be used post hoc to organise event trials into different conditions for comparison in the analysis. For instance, Barch et al (2000) collected overt responses to their verb generation task in order to categorise fMRI trials according to the level of association between each stimulus item and the verb produced by participants. When reporting the study in their paper Barch et al stated that "to our knowledge, this is the first fMRI study to use the content of subjects overt verbal responses as a basis for analyzing the data."

Martin et al (2005) developed a blocked fMRI design to collect overt responses to their picture-naming task. They were successful in their ambitions, however, a block design precludes the categorisation of responses into contrasting conditions necessary for in depth analysis of response types and the associated brain activation patterns. Martin et al admit this limitation with their experimental design and state that "future studies incorporating an event-related design would allow analysis in greater depth." Although they acknowledge that this would be difficult to implement with aphasic patients, due to the often sporadic and unpredictable nature of their spoken output.

A successful attempt to delineate the brain systems responsible for correct and errorful naming performance could have significant implications for the treatment of aphasia. If it can be determined on a case-by-case basis which areas correspond with the production of errors then it paves the way for the use of transcranial magnetic stimulation or training with alternative cognitive strategies in rehabilitative therapies to encourage a more fruitful pattern of cortical activation. It would also promote the use of fMRI as an effective clinical diagnostic tool in speech and language therapy.

1.7 Aims and Hypotheses of the Research Project

The studies designed, implemented and analysed within the research project described in this thesis form an attempt to improve on commonly used methodology to collect more informative data and suggest a new direction for future studies to explore. Research questions posed by the existing body of work in this field were identified and, at the time of beginning this project, were not being addressed, and further, could not easily be addressed using standard techniques. It was thought that a new way of using imaging and behavioural data from aphasic participants was needed to move forward from the stalemate that appeared to exist in the debate of how functional recovery of language does and could occur in stroke-damaged brains. Progression on this issue is crucial for the development of new and effective rehabilitative techniques to improve the quality of life of people with aphasia and their families.

The initial aims of this project were to identify through pilot work a successful fMRI scanning protocol that would allow the continuous acquisition of usable cortical activation data from aphasic stroke patients and, also, to identify the optimal language task to fit the scanning parameters. The scanning protocol needed to allow for the recording of a participant's spoken task responses during the scan session so that it is of a sufficient intelligible quality for the content to be easily understood. The language task needed to promote a mixture of correct and incorrect responses from participants in quantities sufficient for statistical comparison. At all times during this design process a further aim was to take steps to minimise the impact of speech induced motion. It was thought that fulfilling these aims would increase the amount of data available to researchers and enable the combination of simultaneously acquired functional imaging and linguistic behavioural data from each participant so that more penetrating questions could be tackled in this and future research.

The focus of this project was to use the developed imaging protocol and task paradigm in a study with individual aphasic stroke patients who had partially recovered. The aim was to collect correct and error responses to a language task in sufficient numbers to allow for valid statistical comparisons of brain activation patterns involved in each type of response. Considering the results of previous imaging studies of recovered aphasic language, reviewed earlier in this chapter, it was hypothesised that brain areas contributing to successful task performance would be the normal, left lateralised language network that

remained preserved from stroke damage. The hypothesis also states that production of errors would be correlated with an increase of activation in the right hemisphere, homologous to the language areas in the normally dominant left side.

A final objective of this research project was to repeat the study of correct versus error task performance with a group of unimpaired speakers. These participants were given a strict time deadline for responding in order to place load on their language system and promote the production of errors. The aims of this study were to see if cortical activations associated with errors induced from undamaged language networks would be comparable with those from errors arising from a disordered system. It is hypothesised that healthy participants put under additional time pressure would produce significantly more linguistic errors than those without such constraints. Also, as for the patient study, the hypothesis states that error task responses would be associated with an increase in cortical activation within right lateralised areas homologous to the language related areas in the left hemisphere, which would show to be associated with successful responses.

Overall, this project as a whole entity aims to provide a methodological contribution to the general research field of functional imaging of aphasia and, specifically, provide evidence for which brain areas are important for functional recovery of language to take place and, crucially, which areas are implicated in the production of linguistic errors. Insightful results would inform the debate on this topic and suggest a focus for investigations into possible new therapeutic interventions.

1.8 The Scope of this Thesis

This thesis presents the development of an event-related fMRI paradigm using continuous sampling that allows for the collection of responses to an overt object naming task and their post-hoc categorisation into correct and error conditions for analysis.

Chapter 2 reviews the methodological issues that needed to be considered and addressed in the planning of this research. An alternative paradigm is piloted behaviourally in Chapter 3 with healthy participants that utilises the concept of purposefully delaying the naming response. It explores the feasibility of increasing the temporal separation of the cognitive

processing and the mechanics of speech to enhance the usability of data in an equivalent fMRI study.

Chapter 4 details the development of the event-related fMRI paradigm to be used with aphasic participants through pilot work with healthy volunteers. Chapter 5 presents the application of the resulting paradigm in a series of individual case studies with five chronic aphasics. A follow-on behavioural pilot study and subsequent event-related fMRI study with healthy participants is presented in Chapter 6. In these studies participants with normal language function undergo a speeded naming task to induce errors in their performance.

The final chapter discusses all the results obtained in this research and relates them to each other and the literature on recovery from aphasia. Conclusions are made as to how these results contribute to current knowledge in this area and what they suggest for the future direction of research on recovery from post-stroke aphasia.

2 fMRI Issues

2.1 Introduction

Magnetic Resonance Imaging (MRI) of human structures was born in the late 1970s, following a series of discoveries developing the fundamental idea that the behaviour of atomic nuclei could be manipulated by magnetic fields and measured to create spatial images (Huettel, Song and McCarthy, 2004). MRI as a medical application has been in use since the 1980s, but the development of the same principles and technology for studying the brain in action has only recently been realised. The combination of findings of the relationship between neural activity and metabolic changes in the brain and the different magnetic properties of oxygenated and deoxygenated haemoglobin, led to the use of MRI for determining blood-oxygenation-level dependent (BOLD) contrasts.

All of the different functional imaging techniques currently available (i.e. PET, MEG, EEG, TMS) clearly have their advantages and specialties. fMRI, however, has become extremely popular as a non-invasive imager's tool, due to the balance between good levels of temporal and spatial resolution in its images. As a piece of research technology, however, a functional MRI scanner is only as useful as the questions it can answer. So, it has necessarily become very flexible in the types of experiments that can be conducted. New methods of designing, implementing and analysing functional imaging experiments are rapidly being developed to allow the technology to keep pace with the various distributed research fields and their differing needs.

2.2 Aim of the Chapter

Before embarking upon a functional imaging journey it is necessary to understand the technology involved and its strengths and weaknesses relative to the particular field of study it is to be used to explore. It is easy to be seduced by the impressive power of visualising the brain's internal workings, however, its merits should be evaluated to

determine whether this method of investigation is appropriate for the research question, the participants and the experimental procedure to be studied.

It cannot be denied that there are known issues associated with attempting to use fMRI to study typically elderly stroke patients making overt verbal task responses. Before designing the scanning protocol and activation paradigm it is important to review the evidence in the literature to ensure that these issues will not simply preclude any usable and valid data from being collected altogether. Concerns identified early in the experimental process can then be addressed in the study design to limit as much as possible any adverse impact on the data acquired. Therefore, this chapter represents a review of the issues that need to be considered in the context of this project: researching cortical activation patterns of aphasic stroke patients and healthy participants whilst making correct and error responses to an overt language task.

2.3 fMRI in Older Subjects

Depending on the population and the subject being studied, there is a range of challenges to acquiring usable fMRI data. When working within the stroke population it is reasonable to assume that a large portion, if not all, of volunteers will be elderly. There is evidence that the coupling of haemodynamics with neural activity, that BOLD contrasts rely on, may alter throughout a person's life span (D'Esposito et al, 2003). Clearly this would have implications for fMRI studies of neurological conditions often occurring in later life.

Studies which seek to compare groups of healthy, young subjects with elderly, perhaps impaired, subjects on a cognitive task are relying on the assumption that any differences between the groups will be attributable to neural activity. However, this comparison cannot be valid if the underlying vascular workings are not similar for both subject groups.

Several fMRI experiments have been conducted with the aims of testing this assumption. In all studies, groups of healthy young and elderly subjects were compared on the characteristics of their haemodynamic responses (HDRs) and MR signal changes to a simple task. The results lead to some converging conclusions, including the main finding that no fundamental difference was found in the form of the averaged HDR between the age groups (Huettel, Singerman and McCarthy, 2001; D'Esposito et al, 1999). This result

supports the feasibility of researching age-related conditions with fMRI. However, Huettel et al (2001) reported differences in time to peak and return to baseline levels in the older group, coupled with an increase in within subject variability of the HDR shape. This finding suggests that modelling the HDRs for the benefit of statistical analysis may be more problematic in elderly subjects.

The real challenges facing fMRI studies of the elderly population are revealed in the common findings of significantly lower numbers of activated voxels, accompanied by a lower signal to noise ratio (SNR) (see figure 1). D'Esposito et al (1999) note that 25% of their older subject group failed to show any activated voxels within their region of interest, compared to 0% from the younger group. Huettel et al (2001) found that the data from their younger subjects contained on average twice as many activated voxels than their older counterparts. One interpretation of this data is to conclude that the lower SNR in the elderly group is preventing the less strongly activated voxels from reaching threshold levels. This is supported by the fact that voxels showing relatively small effect sizes were less frequently seen in the older subjects.

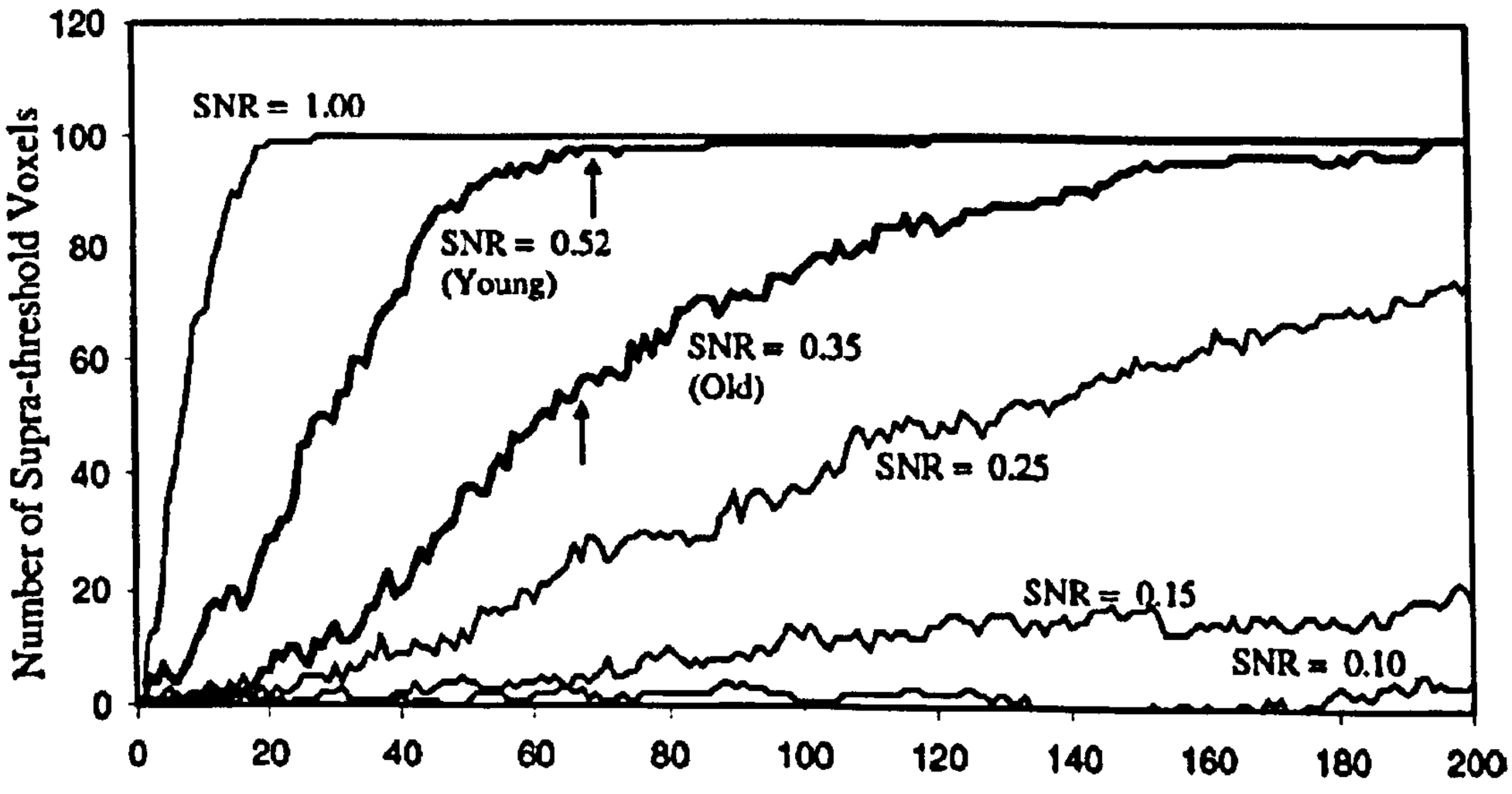

Figure 1. Figure from Huettel et al (2001) representing a simulation of SNRs in young and old subjects across a number of averaged trials, based on data from their experiments.

It is important to note that the lower SNR seen in the older population was due to a significant increase in noise levels, rather than a drop in the magnitude of the detected MR

signal. This means that fMRI is fundamentally still as capable of eliciting and detecting the HDR to a behavioural paradigm in elderly subjects, than in the more routine studies with young people. The decrease in effective HDR detection could be caused by additional physiological noise, e.g. cardiac or respiratory effects, or by a weakening in the coupling of haemodynamic changes and task-related neural activity. The possibility of in increase in head motion being the source of the additional noise was discarded since no relationship was found in any of the studies between the two effects. Huettel et al suggest that studies could potentially compensate for this age-related effect by increasing the number of trials for elderly subjects, thus improving the averaged SNR.

The outcome of these studies leads us to the conclusion that imaging an older population with fMRI is valid and empirically possible. However, the evidence reported suggests that it is more difficult to obtain usable data from these subjects. It also highlights the likelihood that any results acquired may not illustrate the entire story, and represents an additional trade-off between optimal results and the ability to study certain neurological diseases. On a practical level, these difficulties should not be ignored when designing, analysing and interpreting resulting data.

2.4 Imaging the Stroke Population

The issues associated with scanning elderly subjects are compounded when the effects of cerebrovascular disease are present. It should be said that a normal elderly group of subjects may share some characteristics of altered haemodynamics with the stroke population. In addition to changes that may happen in the course of normal ageing, there may be vascular changes present as a precursor to cerebrovascular disease. So it should not be surprising that there is some degree of overlap between the results from the studies discussed above and those seeking to address the feasibility of fMRI in stroke patients.

Concerns with the ability of fMRI to successfully detect neural activation in the presence of this pathology were neatly illustrated in a recent study by Rossini et al (2004). They compared data taken from stroke patients using fMRI and Magnetoencephalography (MEG) techniques whilst doing an identical task. MEG is used to measure magnetic fields evoked by the electrical currents of neurons in the brain. As such, MEG provides more direct, real time information about the activity of the brain than fMRI. Rossini et al used

MEG data as an indicator of real neural activity and compared it to the fMRI BOLD contrast to explore the coupling between the two effects in cerebrovascular patients. It was first established that these two techniques generated the same responses in the control subjects to the activation task. With patients, however, the results showed a divergence between MEG and fMRI elicited activations. All patients demonstrated a recordable electromagnetic response measurable by MEG just as the controls had done, indicating stimulus-related activity in the expected areas in both their affected and unaffected hemispheres. However, the BOLD contrast did not reach above significant threshold levels in the affected hemisphere for half of the patients. In some of these patients, expected areas in their unaffected hemisphere also did not reach significance.

Other studies that compared patients with a compromised vascular system to controls reported similar findings. The magnitudes of the fMRI signal were found to be reduced compared to controls in the presence of comparable levels of neural activity (Murata et al, 2002; Pineiro et al, 2002). Some differences in the temporal form of the HRF were also noted, specifically in the time to peak of the response (Carusone et al, 2002) as seen in the literature on healthy elderly subjects.

2.5 Population Conclusions

Although some aspects of the above literature are quite negative about the ability of fMRI to adequately image cognitive processes in elderly stroke patients, it must be viewed within the context of all other research conducted with this population. fMRI in stroke research is a growth area and a large number of studies that have imaged stroke patients and controls have found similar activation patterns between the two groups of subjects (Cramer, 2004). Differences that have been reported have generally not been huge, indicating that these studies have not suffered from extensive loss of detection power with their stroke subjects, as might be predicted from the above literature.

This literature, however, does invite some caution for studies in this area, but this can be countered by observing some practises to minimise initial variability among participants. For example, effort could be made to select subjects with similar pathologies, or D'Esposito et al (2003) make the suggestion that comparisons could be made between the relative performances on two different tasks rather directly between subject groups.

Concentrating on relative task performance instead of the overall levels of activation between subjects means that differences between groups are more likely to be attributable to differences in neural activity, rather than any confounding effects of vascular changes. In the same way, single case studies where patients are studied in relation to differences in their performance to one or more tasks also benefit from the same kind of internal control.

In summary, there are certain issues that should be taken into consideration when designing and implementing a study using elderly subjects who may or may not suffer from cerebrovascular disease. However, this does not make fMRI studies of this nature an impossible task and there are various experimental techniques that can be used to increase the validity and ease of acquiring valuable results. This will ensure that the advantages of fMRI as a functional imaging method are not lost for this subject population.

2.6 Motion Artifact

2.6.1 Introduction

Noise in fMRI images is something that every imager needs to be aware of, because it could potentially mask the relatively small BOLD signal change and reduce the likelihood of detection. Sources of this unwanted variability can be the functioning of the scanner hardware itself, neural activity not related to the experimental task, differences in task performance and physiological effects, such as heart and respiratory rates. A major contributor to noise in images is physical movement of the head during the scan session. This is problematic because data being collected from specific voxels may actually end up recording the signal from different voxels than intended, as a result of the brain shifting within the field of view. This could lead to areas showing activation when there actually was none of significance (a type 1 error), or areas that were highly activated by the task going undetected (a type 2 error).

A correction is routinely applied to the data prior to statistical analysis to attempt to remedy this. A rigid-body transformation is used to align the series of images using a set of translation and rotation movements. These movement parameters are computed by algorithms that try and produce the best fit between the images needing to be aligned and the images being aligned to. In this way the effects of head motion on datasets can be reduced, however this method does not eliminate all motion-related artifact.

In young, healthy controls the effect of head motion is usually minimal. However, there is evidence that head motion is more prevalent in older subjects and stroke patients have been known to show approximately double the head motion of age-matched control subjects (Seto et al, 2001). This finding obviously gives the issue of movement artifact more emphasis in this field of research.

Most of the sources of variation mentioned occur randomly throughout the experimental time and can be minimised in analysis, when data from all images collected are averaged together. Noise becomes of more concern when its temporal pattern is correlated with the task of interest. It complicates the interpretation of the data when the contributions of true task-related signal and correlated artifact cannot be adequately determined.

2.6.2 Motion in aphasia research

If subjects are required to produce spoken responses to experimental stimuli this is obviously going to produce more head motion than, for example, button presses. In addition to this, the facial, jaw and tongue movements required in speech can create distortions in the magnetic field that are another source of motion artifact in images (Birn et al, 1998).

Traditionally, these issues were neatly sidestepped by researchers conducting experiments into functional and dysfunctional language using fMRI. Participants were instructed simply to think their responses to the task, rather than overtly speaking them. Functional data would then be interpreted as to what the participants' responses should, theoretically, have been (e.g. Perani et al, 2003). Of course, under these circumstances it is impossible to know for certain the true content of the responses, or if the task had been completed correctly or at all. In the case of aphasic participants, who are prone to errors and omissions, it is especially important for the interpretation of results to record the content and quality of responses. Covert speech is an artificial paradigm that truncates the normal progression of language processing by omitting articulation. It is a limiting factor in what aspects of language can be studied, as no information about the quality or characteristics of the task responses can be acquired. There is also evidence from comparison studies that the cognitive processes underlying covert and overt speech may not be the same (Barch et

al, 1998). For instance, Huang et al (2001) found a markedly different pattern of activation when comparing overt and covert generation of speech, notably differential activation in Broca's area. This suggests that the use of covert paradigms may be an unacceptable compromise. With this in mind researchers have spent time developing a number of paradigm design features and methods of results analysis that can work to counteract the effect of these artifacts, making it possible to collect significantly interpretable data from overt activation tasks.

2.6.3 Experimental design

The most significant finding in the search for a viable means of acquiring data in the presence of motion artifact is that the form and temporal characteristics of the artifactual signal change are significantly different from that of the stimulus induced HRF (Huang et al, 2001; Birn, Cox and Bandettini, 2004). Since the BOLD effect is an indirect measure of neural activity its time course is slightly delayed from the point of excitation. Generally, voxels showing true task-related activation will not rise above a significance threshold until at least 4 seconds has passed since stimulus onset (Huang et al, 2001). The HRF peak would then occur 5-7 seconds post-onset. In contrast, signal change induced by motion occurs almost immediately, resulting in a clearly distinguishable temporal pattern than that of the signal change of interest (see figure 2). With this in mind, paradigm designs sensitive to the experimental time course can be used to have a positive effect on the controlling of motion artifact.

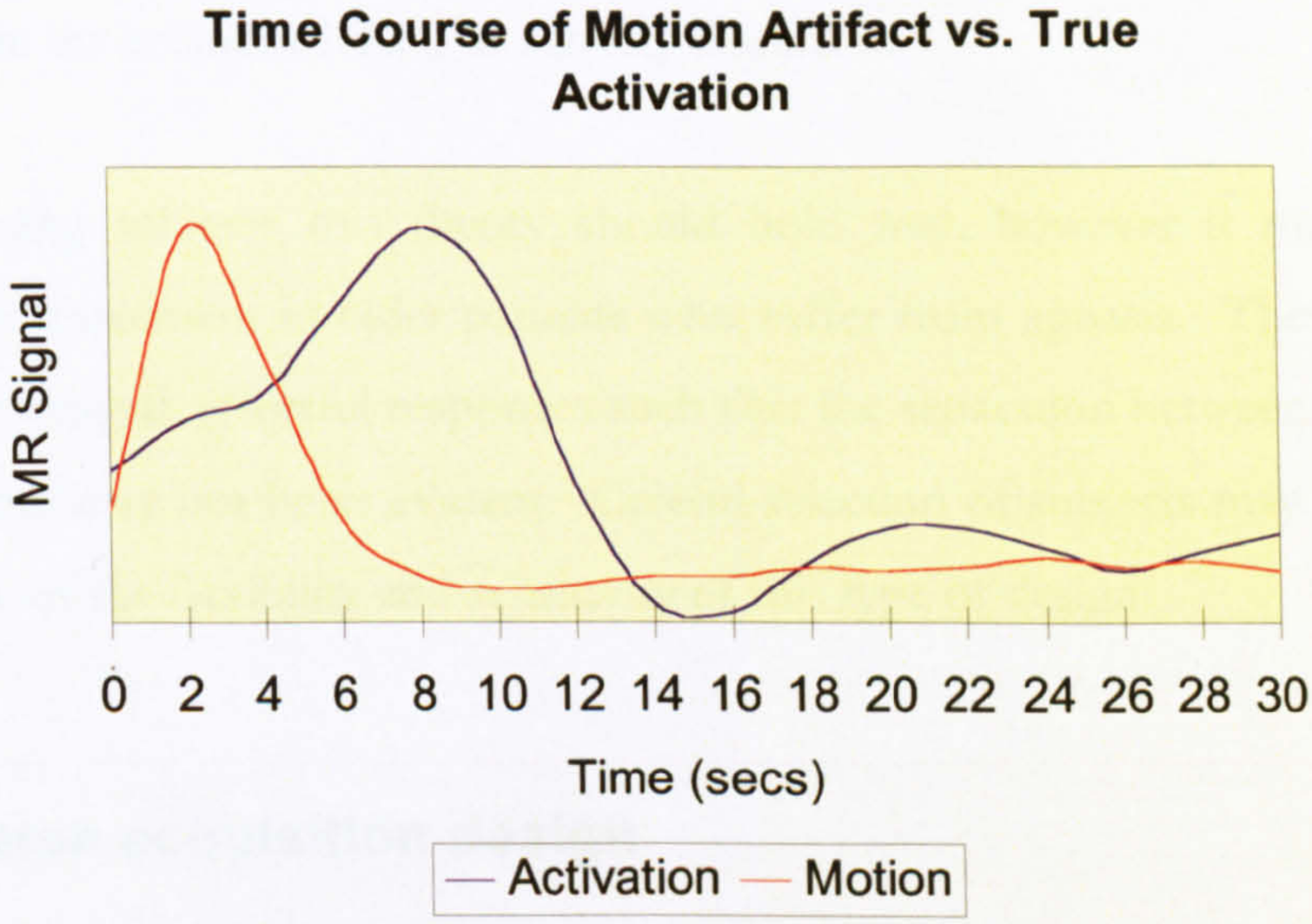

Figure 2. Comparative time courses of the changes in signal related to motion artifact (red) and task-related activation (blue).

2.6.3.1 Event-related design

Early fMRI studies took the lead from the more established PET imagers and concentrated on a study design that organised their task conditions and baseline into on-off blocks. The HRFs from separate stimuli within the task blocks would summate and form a plateau for maximal detection of signal change. Blocked designs can be, therefore, very powerful, but to achieve this it sacrifices sensitivity to the components of the HDR.

Event-related designs are built around a sequence of relatively short stimulus-response parings that are spaced out across the scan session. Individual HRFs can then be averaged across conditions and fitted to a model of the known shape and time course of the canonical HRF to establish where task-related activity occurred within the brain (Buckner, 1998). Time points that fall outside of the model are effectively ignored and it is this feature of event-related protocols that makes it a useful tool to combat early occurring motion-related signal change.

It has been shown that event-related designs are the most effective at detecting activation of interest in the presence of motion-induced activation (Birn, Cox and Bandettini, 2004; Birn et al, 1999; Palmer et al, 2001). The sensitivity of these designs to the temporal dynamics of activity enables them to identify that the motion artifact is not correlated with

the recognisable BOLD response. This identification means that it is much easier to either passively ignore the artifactual data, or actively discard it.

In healthy young subjects this theory should hold true, however it may prove more complicated to implement in older patients who suffer from aphasia. These subjects may make more prolonged, effortful responses such that the separation between the artifact and BOLD response may not be as evident. Careful selection of subjects may be necessary to take advantage of the flexibility and sensitivity of this type of design.

2.6.3.2 Sparse acquisition design

An effective design normally used in auditory experiments is the sparse acquisition technique (Eden et al, 1999). This was developed to tackle the issue of gradient noise generated by the scanner potentially interfering with task-related auditory activations. Using this technique, imaging data is only acquired when the haemodynamic response reaches its predicted peak. At all other times the gradients are switched off (see figure 2). This capitalises on the known delay of the HRF and would mean that any earlier changes in signal would not be collected.

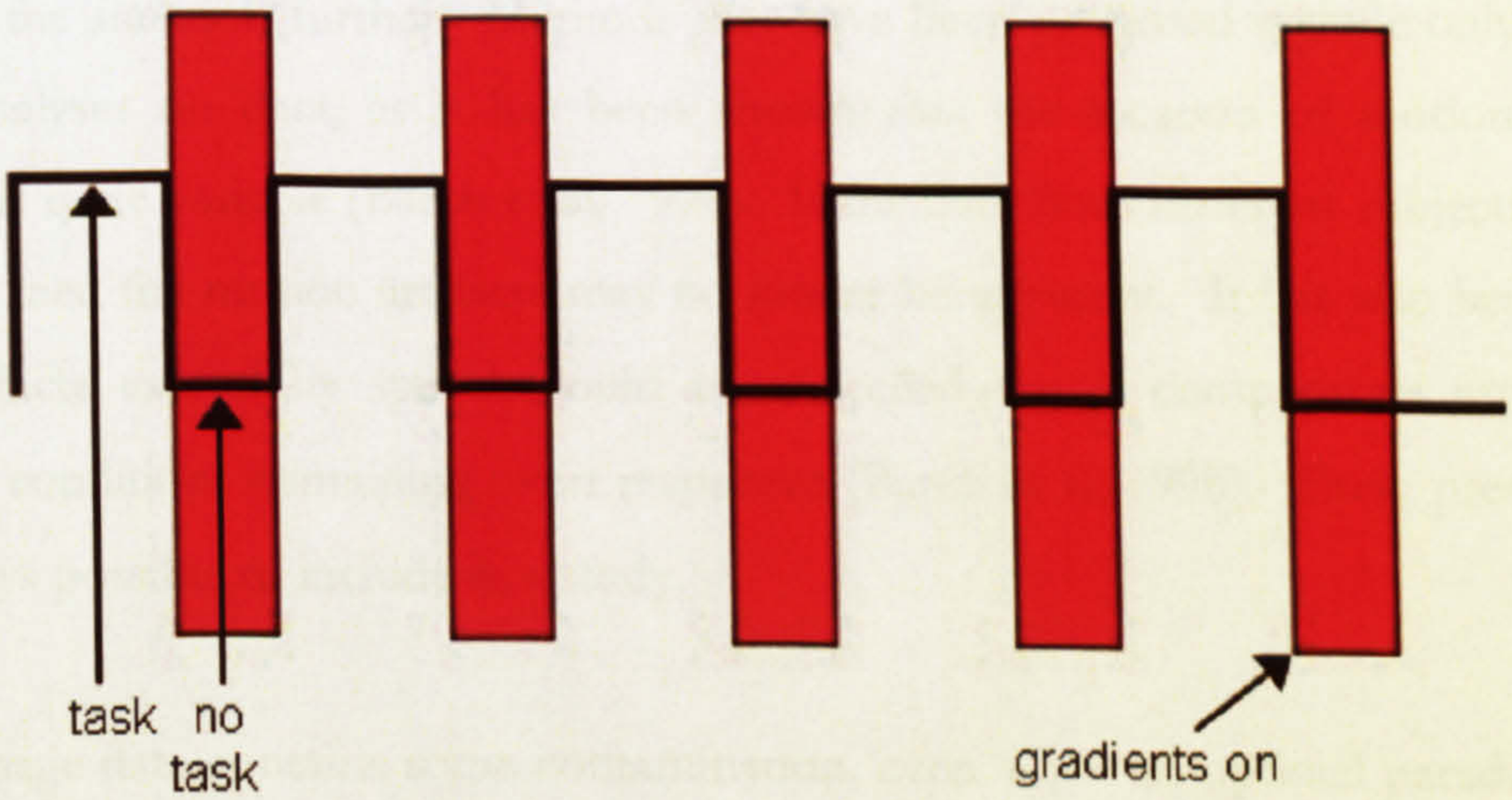

Figure 3. Illustration of the mechanisms of a sparse acquisition paradigm.

The advantages of this method in the context of reducing motion artifact are evident for studies of language. However, its application to studies of aphasia may be problematic. This type of paradigm design depends on the predictability of the timing for the

haemodynamic response peak and this may not be possible to determine in patients whose response times could vary.

2.6.3.3 Delayed response technique

One plausible method for completely separating the artifactual signal change from the effects of interest would be to elicit spoken responses from subjects a specified amount of time after presentation of the stimulus. If the HDR to the task-related neural activity is allowed to peak before the onset of the speech induced head motion, activated voxels can clearly be correlated with the experimental task. This technique effectively amplifies the ability of event-related designs to differentiate between the motion and the true activation.

The viability of this type of paradigm with elderly stroke patients is tested experimentally in the following chapter of this thesis.

2.6.4 Data analysis

Although paradigm design has a substantial impact on the detection power of an experiment, there are additional techniques that can be used with the acquired data to improve the situation further. Methods that have been proposed include only conducting group analyses on data, as it has been shown that the location of motion artifacts in subjects is quite variable (Barch et al, 1998). If the data from different subjects is averaged together then the motion artifacts may no longer be apparent. It has also been suggested that artifacts caused by speech could be cancelled out if comparisons are only made between conditions containing overt responses (Barch et al, 1998). These prerequisites are not always possible to include in a study.

If the image data contains some contamination, even when an optimal paradigm design is used, at the analysis stage any voxels that fall within problematic time points can simply be ignored and not entered into the statistical model (Birn et al, 1999). This has the advantage of totally removing the offending artifact, although necessarily some data will be lost and this may include a portion of the BOLD response if it overlaps with the motion.

A more sensitive technique, and one that has been found to be more successful (Birn et al, 1999) is to explicitly model the time course of the motion-induced signal change as a

regressor. This has the effect of removing artifact without adversely affecting the detection of the HDR.

A combination of one or more of these methods with an appropriate experimental design should ensure the acquisition of detectable and interpretable task-related activity when extra head motion is unavoidable.

2.7 Conclusion

It must be recognised that using fMRI to study elderly stroke patients who have aphasia presents some interesting challenges. However, there is substantial evidence to suggest that it is a viable and valid undertaking and techniques for improving the quality of such data are continually being developed and tested. The ability to conduct flexible, naturalistic and analytical studies, based on the characteristics of the spoken responses, should not be avoided because it is technically more difficult to implement. The benefits of this approach of study to this specific field of research should be pursued to provide insights into the recovery of function within the brain and how this can be enhanced through targeted intervention.

3 Delayed Naming

3.1 Why Explore This?

The risk of data contamination with motion artifact is large when conducting the type of experiment necessary for this research. The inclusion of elderly subjects who have previously suffered a stroke and the elicitation of spoken responses within the scanner pose challenges to experimental design, data processing and analysis (see Chapter 2 for a review). Although there are several techniques available for implementation to reduce the negative effects of head motion on the interpretability of results, their strengths are limited and so cannot provide a guarantee that data can always be rescued.

For instance, researchers have found that an effective method of motion artifact reduction is to employ an event-related experimental design that rapidly samples the BOLD signal at several points across the timecourse of individual stimulus-response pairings (Birn, Cox and Bandettini, 2004). It is known that motion-related signal change displays significantly different spatial and temporal characteristics (Huang, Carr and Cao, 2001) to task-related signal change. Specifically, motion-related signal change reliably occurs 2-3 seconds prior to that evoked by the activation task. This difference cannot be detected using a blocked fMRI design, however, event-related techniques can exploit this separation when modelling effects of interest so that signal changes occurring outside the model are ignored.

Complications with this method can arise when subject's head motion overlaps with their cognitive response to the task. In the scope of this research this could occur when a subject prolongs their overt naming of the stimulus picture. This is not a significant concern for unimpaired participants, who are readily able to follow the instructions to only produce the object name as soon as possible. It was recognised, however, that aphasic patients are likely to experience more difficulty in limiting their responses since they may be more effortful to produce.

People who have aphasia as a result of a stroke and who are able and willing to take part in an fMRI experiment are in limited supply. Since recruitment of patients from this subject group is relatively difficult it is especially important to reduce the risk of discarding data. Therefore, although current methods to improve data quality in the presence of task induced head motion have been found to be largely effective, it was considered prudent to first investigate an alternative method for safely producing a separation of motion and BOLD responses.

3.2 How Does it Work?

It is generally agreed upon that the production of the normal object naming response involves several different processes (Levelt et al, 1998). Briefly, these comprise of initial visual processing, an object recognition system, activation of related semantic concepts, selection of an appropriate lexical entry and corresponding phonology and production of a phonetic plan that is realised through articulation. A delayed naming technique incorporates the same processing stages, but there is an enforced gap between stimulus presentation and the elicitation of the response. Subjects are asked to withhold their naming response until the appearance of a specific cue.

3.2.1 Delayed naming in psycholinguistic studies

Delayed naming has often been used in psycholinguistic studies in behavioural situations to enable researchers to separately evaluate the contribution of the production stage to the overall naming process. Logic dictates that the reaction times from immediate naming reflect the contributions of all processes involved in arriving at a response, whereas delayed naming latencies will only include the articulatory stage. It was introduced in studies as an additional condition to the standard immediate naming paradigm. Forster & Chambers (1973) used the delayed naming task in their study of word frequency effects as a control condition to immediate naming. They needed to ensure that the effects of frequency on naming latencies were not simply due to the intrinsic differences in articulating the names across items. The frequency effect was found not to be present when a delay of 2s between stimulus presentation and response cue was introduced, leading to the conclusion that the effect was not artificially created by variations in item production characteristics.

Since then, the delayed naming technique has been used in many studies alongside immediate naming to try and distinguish the locus of a particular effect within the naming process. If a variable is observed to have an impact on reaction times for immediate naming that is not seen to extend to delayed naming, the logic of this paradigm suggests that the variable must act on processes involved with selection of the appropriate name, rather than the motor response itself. Effects known to have an impact on immediate naming latencies, such as age of acquisition of words, repetition priming and word frequency have been studied under these circumstances (Barry et al, 2001). Delayed naming has also been used with anomic patients who were provided with phonological or semantic cues to aid them with producing the correct name (Best et al, 2002). A response delay was enforced to assess whether the significant influence of cueing on naming accuracy would still exist after a period of time.

In the behavioural linguistics literature the delayed naming technique has been developed to temporally separate different processes in the confrontation naming task. Researchers have set out to distinguish and measure how variables act on different aspects of naming. Is it feasible to employ the same tactic to produce a significant separation of cognitive and motion responses in an fMRI environment?

3.2.2 Delayed response trials in fMRI

The concept of delaying subject responses until the appearance of a cue is known in fMRI paradigms although, perhaps to be expected, the technique has almost been monopolised by studies of working memory. In the working memory literature these types of trials were initially used in studies with non-human primates to locate the neural origins of the encoding, maintenance and retrieval processes concerned. Findings from these studies have been replicated in fMRI experiments with similar delayed response paradigms, implicating a central role for the lateral prefrontal cortex for the delay period of these trials (D'Esposito et al, 2000).

These studies have not investigated language processes, or elicited verbal responses from their participants, but they do emphasise the additional processing involved in a delayed response trial as compared to an immediate response trial. Subjects are required to maintain a representation of the stimulus information across a time period before

constructing the appropriate response. Some studies have also focused upon the fact that the delayed response task has well defined temporal components that can be capitalised on to separately study the different processes at work (Zarahn, 2000). This is the general approach for the current study, that the stimulus presentation to lexical retrieval component can be separately distinguished from the motor response component.

There is one study in the literature that has taken the same approach in their functional imaging study of anomic patients. Cornelissen et al (2003) were interested in tracking cortical changes that coincide with behavioural performance changes in 3 patients who have undergone treatment for their word retrieval deficit. As part of their study they measured brain activation related to a picture-naming task with Magnetoencephalography (MEG). The patients were instructed to wait until a question mark appeared 3 seconds after the removal of the stimulus picture before naming the item aloud. Cornelissen et al made the decision to use a delayed naming paradigm to prevent the muscle movement involved in producing effortful responses from beginning immediately after the stimulus was shown. MEG measurements of brain activity do not incorporate a delay as in fMRI and so it was not necessary for Cornelissen et al to produce a significant temporal separation of movement and task related signal change.

3.2.3 Evaluation of the method

There does exist precedent for the use of the delayed response paradigm in overt language neuroimaging studies and the reasoning for applying it to this research area is sound.

Studies concentrating on the timings of the development of the picture naming response have shown that the neural response to the task develops very quickly after the onset of the stimulus. Levelt et al (1998) conducted a review of studies that had used functional imaging methods with high temporal resolution, such as MEG and Event-Related Potentials (ERPs), to measure the timings of different stages of producing the naming response. With the results they were able to define time windows for the different processes. The initial visual processing and accessing the lexical concept was said to occur 0-150 msec after stimulus presentation, then selection of the lemma at 150-275 msec, phonological encoding at 275-400 msec and finally phonetic and articulatory processing at 400-600 msec. These timings were then supported in Levelt et al's own MEG study, where

they saw peaks of activation following the pattern of well documented naming areas throughout these time windows. Their subjects achieved a mean naming reaction time of 538 msec.

It is important to note that this study and review were conducted with healthy, unimpaired subjects. Aphasic stroke patients are likely to be display longer latencies, especially at the processing stage their deficit is centred on. However, it makes the point that the naming response begins immediately after presentation of the stimulus and develops rapidly from that point. In a delayed naming paradigm in an fMRI study this can be capitalised upon by modelling the changes in signal occurring from the stimulus onset and peaking during the delay period, whilst ignoring the later signal changes that are related to the physical production of the overt response when the cue is presented. This will reduce the likelihood of including artifactual data in the statistical analysis and is the main impetus for investigating this technique.

An additional advantage to the delayed naming paradigm is that subject response times after the cue tend to be quicker than in an immediate naming trial. Cornelissen et al in their MEG study of anomia treatment effects found that in the delayed naming trials their anomic patients all managed to produce their response within 500 msec of the cue being displayed. This is faster than the average response speed in Levelt et al's immediate naming study with unimpaired subjects, described above. This effect of a delayed cue is positive for a study such as the current one, as it is important to prevent patients from prolonging their response. A long response time may interfere with the succeeding trial if sufficient time is not left for levels of signal change to begin returning to baseline after the motion. Also, patients may need time at the end of one trial to prepare for the onset of the next stimulus.

There are a number of disadvantages with moving to a delayed naming task for the current study that must be stated. Allowing activation levels to return to baseline between trials would mean lengthening the cycle time that has been established in the pilot studies. This would either result in less data being acquired or an extension of the overall experiment time. Finding an appropriate balance would need careful consideration as both aspects are of particular concern when imaging aphasic stroke patients.

There is also a concern that patients may make fewer errors in a delayed naming task when they are given significantly longer to prepare their response. The ratio of error to correct responses needs to be kept at a statistically analysable level. It is also recognised that the results from delayed naming tasks in the linguistic literature (mentioned earlier) suggest that effects such as word frequency and age of acquisition that often impact on the difficulty level of the naming task do not seem to extend over a delay period. This may limit the ability to tailor a stimulus set to particular patients to gain an agreeable correct-error response ratio.

What can be drawn from the fact that delayed responses are mostly used for working memory activation tasks in functional imaging is that this technique places more attentional and memory demands on subjects than the standard immediate naming method. Since it is known that elderly participants tend to perform worse in tasks of working memory than younger subjects (Salthouse, 1994) it should be considered that this may affect the ability of the stroke patients to adequately perform the task. Also, working memory may have been affected by the patient's stroke and therefore could be a contributing factor in their aphasic condition.

These issues need to be taken into consideration when trialling the delayed naming method with aphasic patients. However, it can be said that for an fMRI study the delayed naming task does represent the advantages of the popular covert naming paradigm without the disadvantages surrounding not recording the task responses. The study presented here is a behavioural pilot study, conducted outside of the scanner environment, to test the feasibility of using the delayed naming paradigm with aphasic patients. This paradigm will also be compared with the same patients' performance on the standard immediate naming task.

3.3 Aims and Hypotheses

The aim of carrying out this pilot study of delayed naming is to trial the methodology with aphasic stroke patients who have word finding deficits, to see if correct and error responses to a confrontational naming task are able to be collected in this way. This technique of temporally separating the stimulus presentation from the problematic verbal response could potentially strengthen the SNR significantly for sampling of the cognitive processing

taking place. This would increase the statistical power of the data collected and thus improve the ability to make inferences relating to which brain areas are important for successful naming performance and which are involved in the production of errors.

Before this can be tested within the fMRI environment, however, it is important to determine whether requesting aphasic patients to withhold articulation until a cue is given is an appropriate or even feasible way of extracting true task responses. For the purpose of the main fMRI patient study it is critical that delaying articulation for several seconds does not change the quality or content of the response altogether from that which would have been produced immediately. Another result to be checked in this pilot study is the ratio of correct and error trials produced via this experimental method. Relative numbers of each type of trial need to be enough for statistical power to be at a level where valid statistical comparisons can be made.

Hypotheses to be tested in the delayed naming pilot study are that there will be no significant difference in the number of correct or error trials between the immediate and delayed naming conditions, but that response times will be shorter when naming is delayed than when it is required immediately after stimulus presentation.

3.4 Experimental Procedure

3.4.1 Patients

Six aphasic stroke patients were recruited for this study, three of whom were also tested in the fMRI study of aphasic patients and three who had been excluded from the scanning study. All gave informed consent and the Nottingham City Hospital ethics committee gave ethical approval for the study.

Prior to the behavioural study all subjects were given a series of assessments relating to their ability to perform the task as instructed. These included an object perception task from the Visual Object and Space Perception battery (VOSP), tests of verbal and nonverbal agility taken from the Boston Diagnostic Aphasia Examination (BDAE), parts 1 and 5 of the Token Test and the Pyramids and Palm trees assessment of semantic processing. These tasks probed the patients' visual perception, language comprehension and verbal memory, semantic associations and tested for the presence of dysarthria. The

Graded Naming Test was also administered to gain a measurement of patients' naming ability independently of the demands of the experimental task. The assessment scores for each patient are summarised in table 1.

Patients	Object Decision	Nonverbal Agility	Verbal Agility	Token Test (1)	Token Test (5)	Pyramids & Palms	Graded Naming
1	20	6	6	10	15	42	4
2	19	1	7	10	8	51	13
3	20	6	6	10	1	50	12
4	15	4	6	10	0	48	13
5	15	4	-	2	0	48	0
6	13	3	-	9	0	40	0

Table 1. Showing assessment scores from all 6 patients. Maximum possible scores are: object decision 20, nonverbal agility 12, verbal agility 14, Token Test (1) 10, Token Test (5) 22, Pyramids & Palm Trees 52, Graded Naming 30. Missing scores from the verbal agility test mean that the patient missed 2 items or more and so invalidated the test.

Only one patient fell below the cut-off score of 14 for the object decision task, indicating that visual perception of objects was not a significant problem for this patient group. All patients showed some degree of impairment on the tests of verbal and nonverbal agility, which are sensitive to weakness in articulation and muscle movement. Patients 5 and 6 showed the greatest difficulty with these tasks. The Token Test part 1 was completed successfully by most patients, with patient 5 being the most impaired. Most patients struggled with part 5 of the Token Test, with only patient 1 attempting more than half of the test. Other patients mixed colours, shapes and types of command (i.e. putting a token in front of another when instructed to place one under the other), or only carried out part of the instruction. This indicates that this group of patients may have difficulties with auditory comprehension and working memory span. The Pyramids and Palm Trees assessment was reasonably well completed, with patients 1 and 6 showing a mild impairment on this semantic test. The Graded Naming test proved difficult for most patients, giving an indication of the word frequency level that begins to cause problems.

3.4.2 Method

The study consisted of two conditions, immediate and delayed naming, which were presented in 4 blocks using an ABAB configuration. The order of the blocks was counterbalanced across patients to eliminate any confounding effects of tiredness. There were 30 trials in each condition, so 60 15s trials in the experiment in its entirety. The immediate naming condition mimicked that used in the fMRI pilot studies. Stimulus presentation duration was 1s with a 14s interstimulus interval (ISI) and patients were instructed to produce a single response as soon as possible. The delayed naming paradigm was designed to be as close in timings to the immediate naming condition and fMRI pilots as possible (see figure 2). Trials were again 15s in length, with a 1s stimulus duration. At 7s into the trial the response cue was displayed for 1s, at which the instruction was to produce a response as soon as possible. In both conditions a fixation cross was present in the centre of the screen at all times when stimulus or response cue were not displayed.

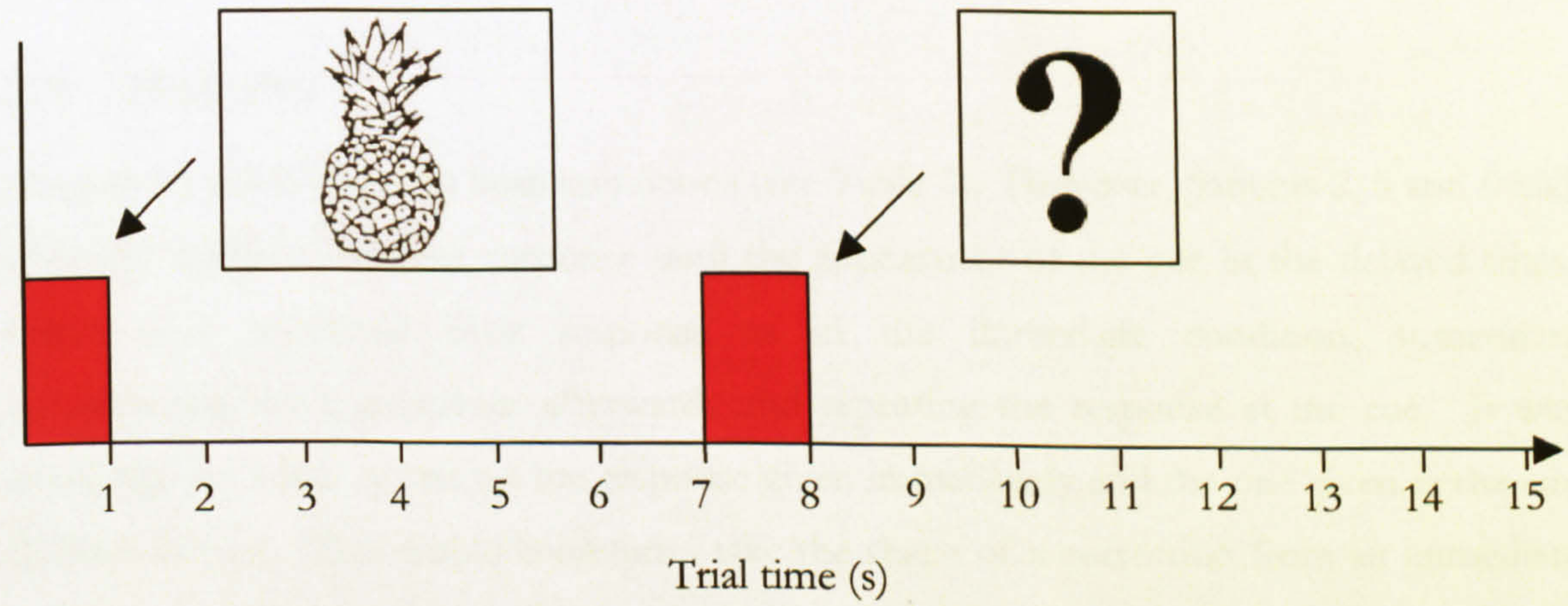

Figure 2. Timings of trial in the delayed naming condition.

30 black and white line drawings were selected from the Snodgrass and Vanderwart (1980) standardised stimulus set, also used in the fMRI pilot studies. An equal number of pictures were selected from the 10 major semantic categories in the set. Pictures were selected that had the lowest word frequency scores according to figures from Snodgrass & Vanderwart. This was done to maximise the ratio of errors to correct responses, especially in recognition that patients may show a reduction in errors in the delayed naming condition due to an increase in response preparation time. The same stimulus set was used in both conditions, for ultimate control of item characteristics between the immediate and delayed trials. The stimuli were randomised within each condition for each patient. This coupled

with the counterbalancing of trial blocks meant that the condition that a stimulus item appeared in first was also counterbalanced across patients.

Responses were categorised as correct when they matched the item name or were a common synonym. Errors were identified as any given response that differed from correct. These included errors semantically related to the target name, formal errors that shared some phonemes with the target but were not semantically related words, mixed errors that shared phonemes and a semantic relationship, words unrelated to the target in any way and non-words that did not conform to a word found in an English dictionary. Errors were not formally broken down into these sub-categories as this study is concerned with the overall number of errors produced by patients. Responses were classified as no responses when the patient made no attempt to produce a name, or when they indicated that they were unable to attempt to name the stimulus. When patients gave two responses, against the instructions, only the first complete response was included.

3.5 Results

All patients made errors in both conditions (see Table 2). However, patients 2, 5 and 6 had difficulty withholding their response until the appearance of the cue in the delayed trials. Often they produced their response as in the immediate condition, sometimes remembering the instructions afterwards and repeating the response at the cue. It was noted that on some occasions the response given immediately and the one given at the cue differed in type. This would sometimes take the shape of a correction from an immediate error (e.g. stimulus = axe, response = hammer > axe), or producing an error at the cue when previously the patient had answered correctly (e.g. stimulus = harp, response = harp > axe). Anecdotally, it was also evident that patients 2 and 5 would sometimes engage in whispered rehearsal of their response in the delayed naming trials until the cue was presented.

Patients	Immediate Naming			Delayed Naming		
	Correct	Error	No Response	Correct	Error	No Response
1	25	5	0	24	5	1
2	18	12	0	21	6	3
3	21	2	7	23	1	6
4	23	6	1	25	4	1
5	8	13	9	9	5	16
6	21	8	1	18	11	1

Table 2. A breakdown of response types produced by patients in the immediate and delayed naming conditions.

For those patients (1, 3 and 4) who completed the delayed naming study, according to the instructions of the task, it was found that levels of response types did not differ significantly in the two conditions. Patients 1 and 3 achieved faster response times in the delayed naming condition, with mean decreases of 650 msec and 620 msec respectively (see figure 3). However, patient 4 means showed a 310 msec increase in the response time for the delayed naming condition.

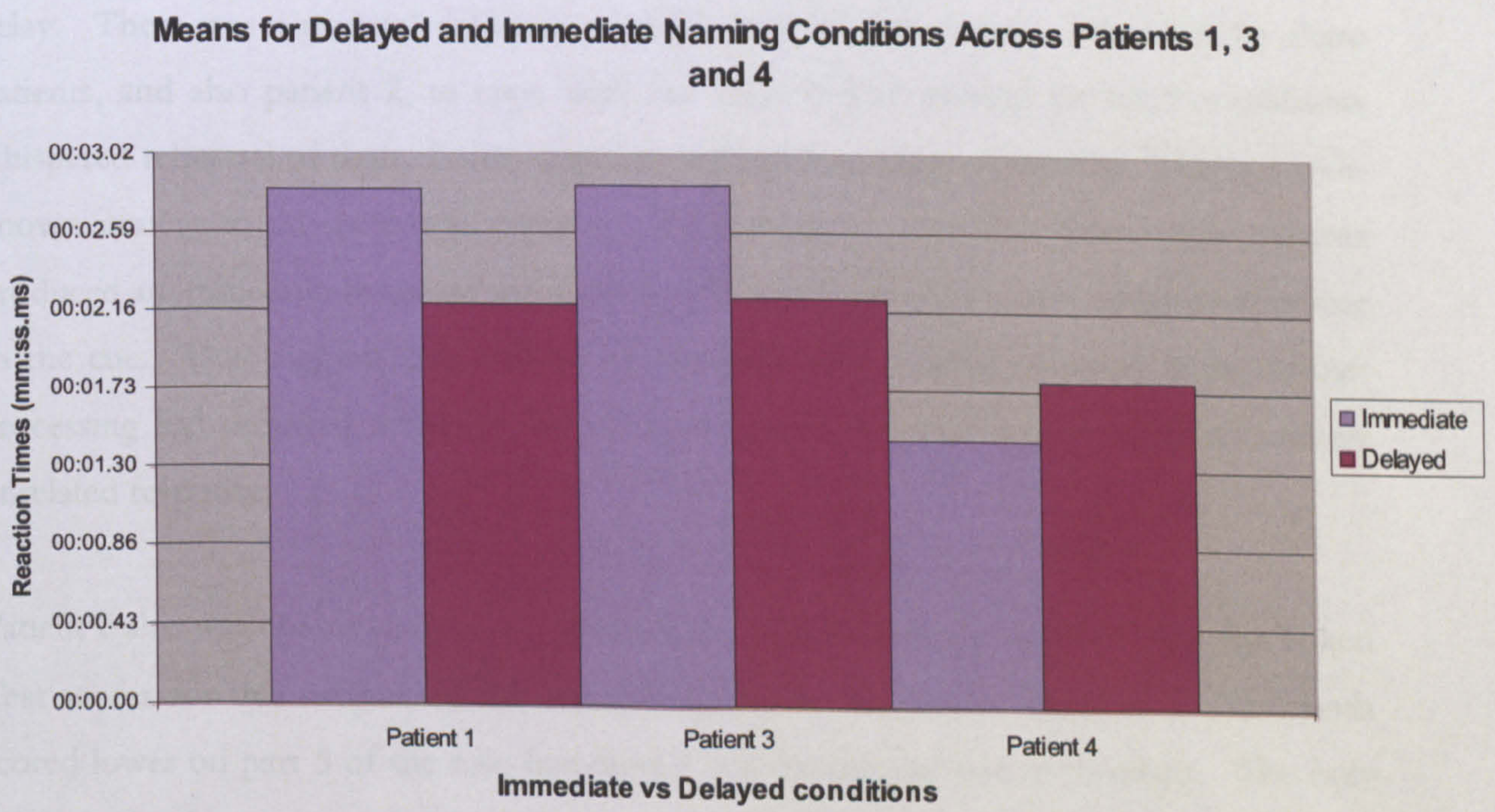

Figure 3. Mean response times for patients 1, 3 and 4 for the immediate and delayed conditions.

3.6 Discussion

The patients who performed the delayed naming task correctly are also the patients who made the least amount of errors and the largest number of correct responses in both conditions. This is a problem in a research project that is concerned with comparing the two response types in a statistically meaningful way. From the outcome of this pilot study it seems that patients whose naming deficit is large enough to be of benefit to the main functional imaging study are not able to carry out the delayed naming paradigm satisfactorily. From the initial assessments it can be seen that patients 5 and 6 are generally more impaired across the tests than other patients. They produced low scores in the object decision, verbal agility, Token Tests and the Graded Naming test. Deflated scores in the assessments could implicate a number of reasons why these patients had difficulty with the study demands.

It was thought likely that low scores in the Token Test would predict which patients would struggle with the delayed naming task. The Token Test is designed to look at aspects of language comprehension and short-term memory of information that the subject is asked to act upon. Patients who have difficulty with this assessment may not have fully comprehended the instructions for the study, or they may have trouble with retaining the stimulus information or outcomes from their own internal processing across the required delay. There was anecdotal evidence of difficulties of this nature. Attempts by these patients, and also patient 2, to cope with the delay before naming included continuous whispered rehearsal of their chosen response until presentation of the cue. This is a well-known strategy to aid short-term memory. There were also the trials where these patients produced an immediate response and then an additional but significantly different response to the cue. This suggests that instead of retaining their original response some further processing had occurred, either to correct a recognised mistake or to produce an entirely unrelated response.

Patient 2 also was unable to correctly perform the delayed naming task, however the Token Test scores for this subject are the second highest of the group. Patients 3 and 4 both scored lower on part 5 of the test, but carried out the delayed task accordingly. The only assessment where patient 2's scores were lower than the rest of the group was nonverbal agility. It is unlikely that a deficit in this test would cause difficulty with the delayed condition without similarly affecting the immediate naming. It is possible that this patient

misunderstood the task instructions, or has more attentional problems that were not picked up by the assessments. It should be noted though that patients 3 and 4 performed similarly on the Token Test parts than patient 6 without displaying any difficulties with the delayed naming task. It would seem that the assessments carried out prior to this study are not sensitive to the ability to properly execute the task.

An additional issue to consider is that erroneous delayed naming trials, such as those described above, would prove costly in an fMRI study. Constant whispering would increase head motion and artifacts associated with movements of the tongue. It would be very difficult to disentangle the task-related signal if motion artifacts overlapped the entire trial time. Also, if the response given at the delayed cue were to be different from that prepared in earlier processing it would render use of the delayed paradigm useless. Trials may be included in one condition when the neural activity being measured truly reflects that of a different condition.

Conducting this pilot study has highlighted several flaws with the use of the delayed naming paradigm with aphasic patients who are significantly impaired.

3.7 Conclusion

The central issue with the results from this pilot study is that patients who have enough of a naming deficit to generate a correct-error response ratio nearer to the optimum are not able to perform the delayed naming task properly. This obviously would pose a significant problem to the research project if this paradigm was adopted. It was judged that this problem outweighed the advantage gained by switching to this paradigm, of creating a larger temporal separation between task-related activity and motion-induced artifact, and would confound any conclusions formed regarding the brain areas implicated in the recovery of functional language in the presence of stroke damage. Therefore, on the basis of this pilot study it was decided to select the more commonly used immediate naming paradigm for use in the main patient fMRI study of correct versus error naming responses.

4 Paradigm Development

4.1 Introduction

In fMRI studies there are a large number of variables that can be manipulated to produce differing results for the same experimental paradigm. Scanners are expensive to run and for projects such as this, where the numbers of potentially viable subjects are limited, it is vital to acquire statistically usable data for every scan session. These issues serve as motivation for spending a significant amount of time researching the optimal paradigm design and scanning decisions for a particular experiment. Each decision will have an impact on the statistical power to detect "real" neural activity and, therefore, will affect the interpretability of the data and the overall success of the experiment.

The primary purpose of this research project was to contribute to the understanding of the mechanisms involved in recovery from aphasia and the relative roles each hemisphere takes in reinstating language function. This was to be achieved by exploring aspects of the production of errors and normal responses in aphasic patients and unimpaired subjects when undergoing a simple object naming task. Therefore, the methodological challenge was to develop an experimental paradigm and imaging protocol that would produce interpretable datasets for both populations under these conditions. This experimental design would then need to be piloted successfully before embarking upon the study of successful and unsuccessful name retrieval from impaired and unimpaired language systems.

4.2 Event-Related Issues

The nature of the experimental task demanded that an event-related paradigm be used for the fMRI studies conducted as part of this research. It is not possible to know apriori what response a person with word finding difficulties will make to a given set of stimuli. Studies concerning the consistency of responses made by aphasics to naming tasks have shown that although they may demonstrate a tendency to produce a similar overall proportion of

different types of responses to repeated presentations of picture stimuli, internal consistency between specific items can be low.

Howard et al (1984) demonstrated this in their study with 12 aphasic patients of different diagnostic types. Patients were twice presented with a set of 300 line drawings of objects that they were asked to name. It was found that patients often produced an identical response to the same items in both presentations, although this proportion did vary across patients from 35% to 74%. However, when correct responses and omissions were taken out of consideration so that only the proportions of identical errors were examined, the percentage of response consistency dropped to between 5% and 43%. If a patient had responded correctly to the first presentation of the stimulus then results showed a correct response for the second presentation was significantly more probable. An error on the first presentation did not make any class of response statistically more likely in a further test.

Similar results to this have been reported in subsequent studies (Lambon-Ralph, 1998). This means that stimuli cannot reliably be blocked into different conditions, i.e. correct and error response conditions. Since event-related designs allow for the measurement of individual BOLD responses to stimuli, events can be sorted into their respective conditions post-scan session based on the characteristics of the subject's responses to the task.

The flexibility of the type of experiments that can be explored using event-related scanning methods is a big motivation for pursuing this technique. However, there are other positive and negative aspects to this design choice that must be taken into consideration when deciding upon the parameters of the experiment.

4.2.1 Detection power versus estimation efficiency

In general, there are two central issues researchers may be concerned with when designing fMRI experiments. Will the experiment have enough power to detect the underlying task-related neural activity, and will the design provide the ability to accurately estimate the characteristics of the measured HDR. It has been shown, however, that these two goals cannot be maximally achieved within the same experimental design (Birn, Cox and Bandettini, 2002; Liu et al, 2001).

Blocked designs consist of long alternating periods of stimulation and rest, resulting in a large change in signal between task and baseline conditions that can be more easily detected in analysis. In order to achieve this contrast the signal from repeated task stimulation is allowed to summate over the block time. This has the effect of masking the true shape of the HDRs in return for higher signal amplitude. Therefore, although this type of design is preferred for detecting significant levels of activation, it is poor for sampling the HDR characteristics. The trade-off is reversed for event-related designs where individual trials are randomly intermixed. Estimation of the HDR is more optimal, but with an accompanying reduction in detection ability.

Estimation efficiency is beneficial when an aim of the study is to focus on the timings of neural activity or to separate out different processes that may be occurring within a time period. However, detection of neural activity must be attained for any inferences to be made about task-related processing and so a major consideration of event-related designs must be to achieve a level of statistical power where interpretable results are possible.

4.2.2 Interstimulus interval timings

One simple method to increase the ability of an event-related design to detect significantly activated voxels in images is to increase the number of events in the scan session. When scanning subjects from populations such as stroke patients it is not desirable to run experiments for a lengthy amount of time. These subjects can tire easily and disengage from the experimental task or become uncomfortable and increase their levels of movement within the scanner. A potential workaround for this issue would be to decrease the interval between stimulus presentations in order to use the experiment run time most efficiently.

Fast running designs do not allow the HDR to return to baseline before delivering subsequent excitations and so for analysis they rely on the ability to extract the relative signal contributions of individual responses. This assumes that the response to each stimulation will be identical and that they will add together in a linear manner. There is some evidence that this is can be roughly true (Dale and Buckner, 1997), however, some clear non-linearities have been found in results that complicate the analysis, relating to the influence of preceding trials on the HDR (Friston et al, 1998; Bandettini and Cox, 2000).

Trials arranged close together in time can suffer from attenuation of the signal amplitude cause by saturation of the signal. This can impact on detection of activation and the ability to accurately model the HDRs in the data. These problems can especially be seen at short interstimulus intervals (ISIs) of >6 seconds, showing that HDRs that are not sufficiently spaced throughout the experiment are not independent.

Obviously, in an overt language paradigm individual trials need to be long enough to allow for a response to be given. This is even more pertinent with subjects whose language processing is disturbed and are likely to display slower reaction times. For this reason it is not practical for these patients to be subjected to rapidly occurring trials. Also, work by Bandettini & Cox (2000) exploring the optimal ISI to be used in event-related designs, when the ISI is to be held constant, suggested that for a stimulus duration of 2 seconds or less an ISI of 12 seconds produced the largest signal change (see figure 1). This allows the elicited HDR to return to baseline prior to the next excitation. It also allows for sampling of the baseline levels for use in task versus rest results contrasts.

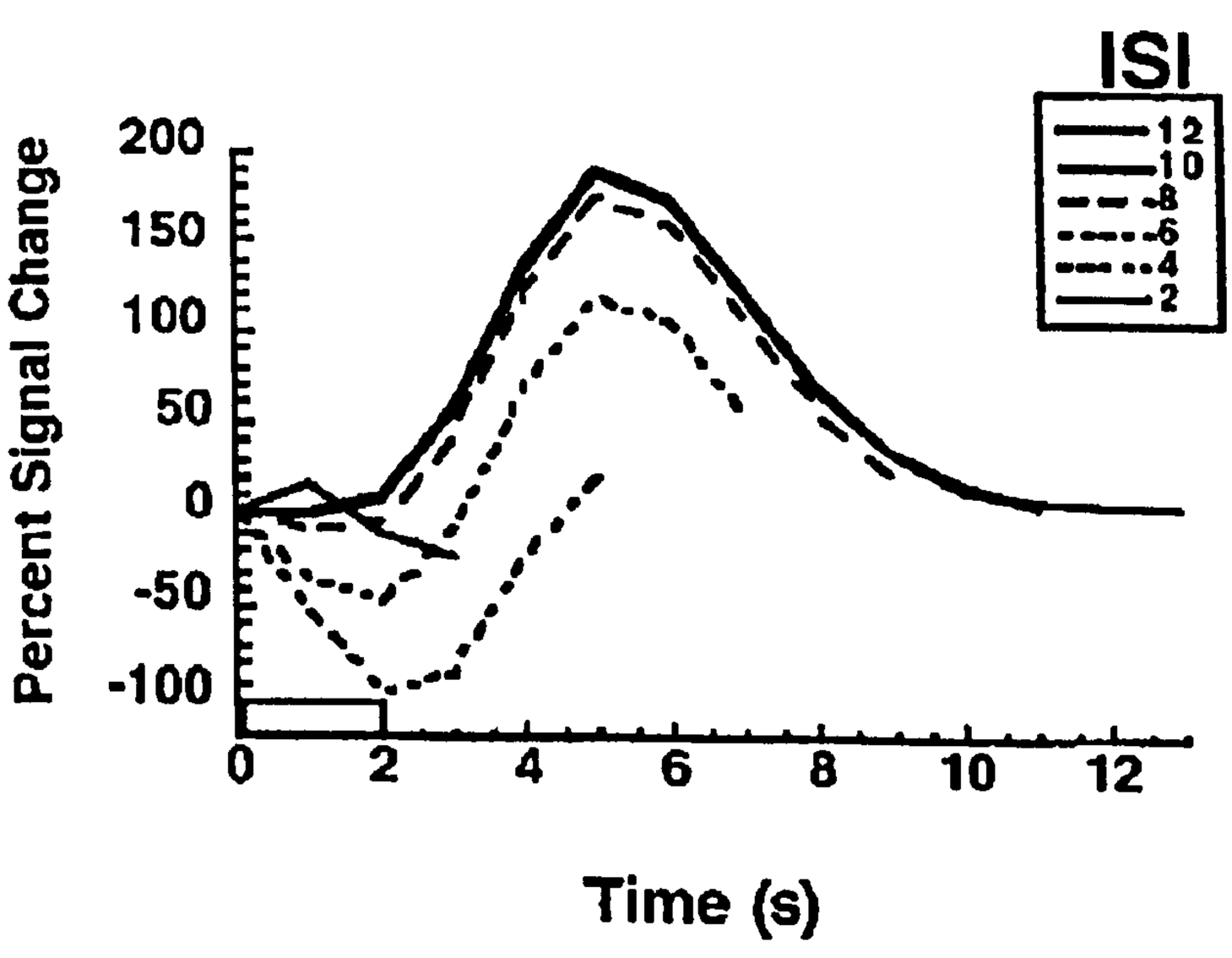

Figure 1. Taken from Bandettini & Cox (2000) showing the differences in percentage signal change between event-related data acquired with varying ISIs.

Taking all the above evidence into consideration it was deemed preferable for this project to opt for a design with longer ISIs, to aid patient co-operation, eliminate the chance of linearity errors and to permit sampling of the entirety of the HDR.

In order to assess the success of the chosen paradigm and the design decisions made the experiment was piloted with healthy volunteers before aphasic stroke patients were recruited. Success would be judged by comparing the activation patterns with those from previous fMRI studies of confrontation naming.

4.3 Aims and Hypotheses of Pilot Work

The principal aim of the pilot work was to test an event-related BOLD fMRI scanning protocol with unimpaired participants whilst engaged in a confrontation naming task requiring verbalisation of responses. The piloting needed to determine whether the scanning parameters and activation paradigm were sufficient for the acquisition of interpretable fMRI datasets and for successful completion of the behavioural task. All equipment to be used for the task had to be tested within the environment of the 3T magnet, such as the computer and screen needed for stimulus presentation and the microphone setup for the recording of participants' responses. The sound recordings containing the responses needed to be assessed for quality and clarity. An additional aim of this testing period was to ensure that participants were comfortable and happy with the experimental scenario and that the demands of the task and length of time within the scanner's confines did not have a negative impact on them. It was intended that through pilot work an imaging protocol and task paradigm would be developed that could be applied with aphasic patients to explore the recovery of functional language after stroke damage, by identifying brain regions involved in successful and unsuccessful performance on a confrontation naming task.

The specific hypothesis to be tested during piloting was that functional imaging data would show increases in signal in areas known to be involved in object naming processing when picture stimuli is presented to the participant for naming. These areas may include primary visual areas in the occipital cortex, posterior inferior frontal cortex, anterior, posterior and inferior areas of the middle and superior temporal cortex, inferior parietal cortex and cingulate cortex (DeLeon et al, 2007). It is expected that activation will be seen bilaterally in these areas with a tendancy for dominance in the left hemisphere. Statistical evidence to support this hypothesis would serve as validation for the study design for use in

investigating the cortical areas underlying successful and unsuccessful word retrieval in aphasia and normal language processing.

4.4 Pilot Study 1

4.4.1 Subjects

Three subjects participated in this initial pilot study, all male with an age range of 20 to 40 years, mean age 32. Subjects were healthy volunteers who had normal language ability, with English as their first language, and no history of neurological disease or damage.

4.4.2 Paradigm

For this to be a valid pilot for the study with aphasic patients it should use the same activation task and paradigm. The activation task the subjects were required to perform within the scanner was a simple naming to pictures task. This task was chosen for the patient study, because naming appears to be the linguistic ability that is most consistently impaired across different aphasia types, with many patients producing errors and omissions when attempting successful object naming (Howard and Gatehouse, 2006). Therefore it is the task that most aphasic patients are likely to produce a ratio of correct and error responses for. Confrontation naming is a task often used in functional imaging studies of language and so the normal neural correlates of the task are fairly well defined by previous work. Also, a picture naming task is relatively easy to set up, and for subjects to perform, within a scanning environment.

Subjects were shown line drawings of everyday items and were asked to name the item aloud as soon as they could. The pictures were presented to them on a screen placed just outside the scanner. Subjects were able to view the screen from their position on the scanner bed with prism glasses, which redirected their vision outside the bore of the scanner.

Their responses were recorded digitally using a small microphone connected to a laptop computer. The microphone is encased in a plastic tube and rests approximately a metre from the scanner. The tube extends into the scanner where it was attached to a plastic funnel fixed underneath the subject's chin. This funnel allows responses to be directed

down the tube to be picked up by the microphone whilst the scanner is running, without any need for subjects to move their head unnecessarily when speaking. This microphone setup was based on the system used by Barch et al (1999) in their studies of the efficacy of capturing spoken language within an fMRI experiment.

This task should evoke the stages of cognitive processing outlined within figure 2. When the stimulus picture first appears on the screen subjects should be engaged in visual object identification, activating their primary visual cortex within the occipital lobe, and processing stimulus features and form, with information flowing through a ventral pathway projecting to the temporal lobe. Once the object has been recognised visually there will be an attempt to link it to a particular semantic concept, revealing its functions, semantic properties and place within the subject's world. This is thought to occur across various areas in the temporal lobe, such as the fusiform and inferior gyri down to the temporal pole (Damasio et al, 1996). A noun will be selected that describes the concept and information about that specific word's lexical form and component speech sounds will be accessed ready for articulation.

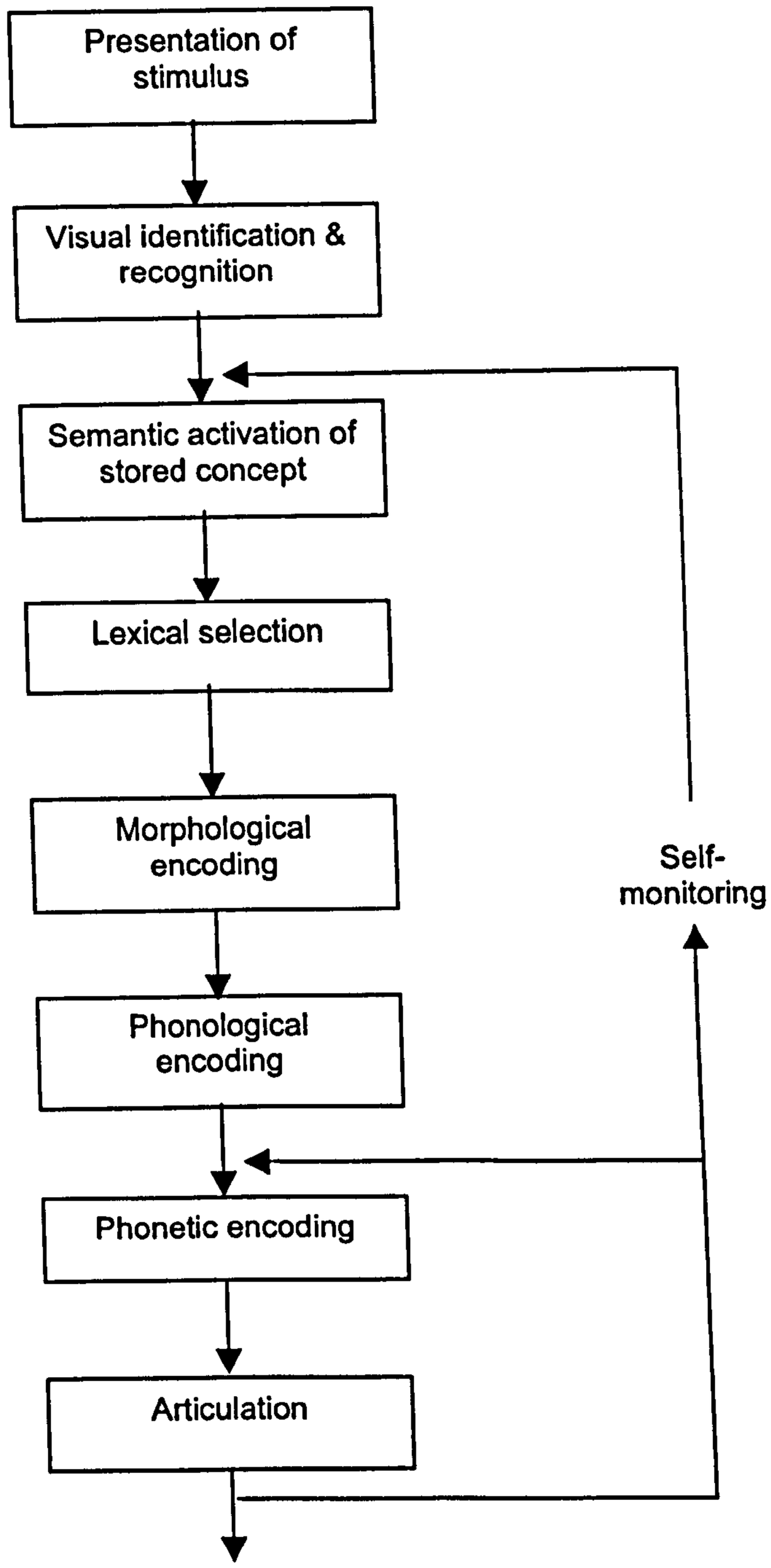

Figure 2. A generalised box and arrow model of the object naming process.

Within the paradigm, stimulus pictures were presented for the duration of one second in every scan cycle. This value was chosen after it became apparent during tests that the longer the picture presentation time, the slower the response times. Subjects then viewed a baseline fixation cross placed in the centre of the screen for fourteen seconds. This duration is to allow the haemodynamic response in the brain elicited by the task to decay and return to approximately baseline levels ready for the next excitation. Each cycle was

therefore fifteen seconds long and was repeated sixty times, giving the paradigm a total running time of approximately fifteen minutes (see figure 3).

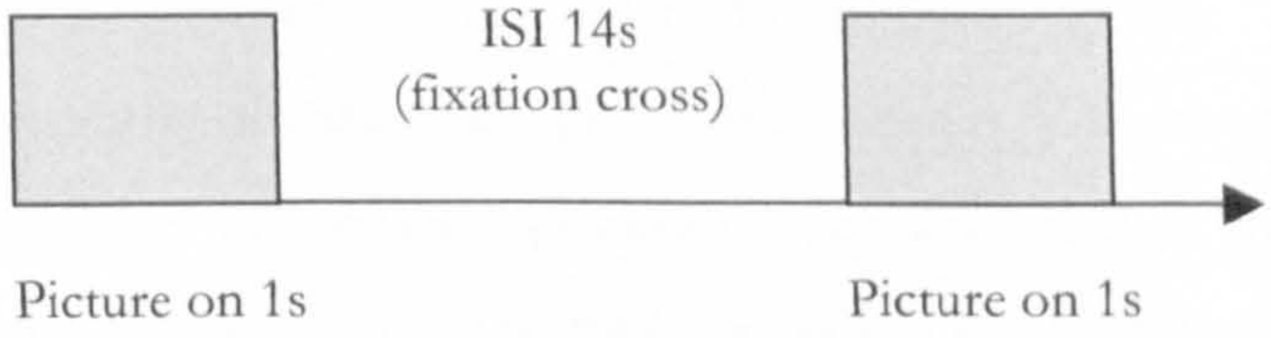

Figure 3. fMRI paradigm timings for both the pilot study and the subsequent patient study.

4.4.3 Stimuli

All 60 stimuli used in the activation task have been selected from the standardised set of pictures published by Snodgrass and Vanderwart (1980). These are black and white line drawings that have been especially developed for use as stimuli in psychological experiments. The pictures have been standardised for their visual complexity and familiarity and are presented as exemplars from a number of different semantic categories.

For the purpose of this project 60 pictures from the 10 major categories were selected, six pictures from each category. The selection was made on the basis of the item's familiarity to the average person living in the Midlands in England. Figure 4 gives examples of the style of picture used.

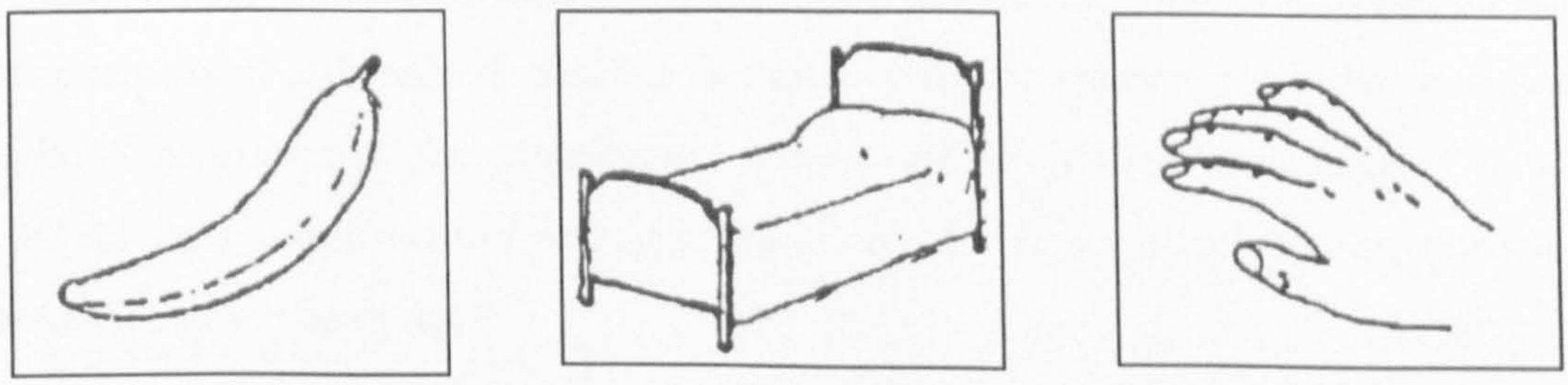

Figure 4. Some examples of the stimulus set used, from Snodgrass and Vanderwart (1980).

4.4.4 fMRI protocol

Functional T2* weighted images were acquired coronally using echo planar imaging (EPI) at the strength of 3.0 tesla (3T) using an fMRI scanner built in-house at the Sir Peter Mansfield magnetic resonance centre at the University of Nottingham. Twenty-two slices were collected in each volume of data to cover the whole brain. The image matrix was 64 x 64, with an inplane resolution of 3mm x 3mm and a slice thickness of 9mm. 300 volumes were acquired during the scan session with a volume repetition time of 2.992s (TR=2.992s). The first four volumes were discarded to allow the MR signal to stabilise, leaving 296 volumes to be analysed further.

T1 weighted images were acquired at the end of each subject's scan session to provide structural information for later use in interpretation of functional analysis results.

4.4.5 Data processing and analysis

Sound files containing the digital recordings of the naming responses were passed through a scanner noise cancellation tool developed by Cusack et al (2005). This removed excess scanner noise from the files increasing the clarity of responses for transcription. Participants were not expected to produce many errors therefore the transcriptions were simply used to ensure successful task completion.

Before statistical analysis is applied to the data functional and anatomical images need to be put through a series of preparatory stages. These pre-processing steps are primarily designed to reduce potential sources of error in the data and enhance the ability to detect the relatively small task related changes in signal from the general noise that is apparent throughout the course of the experiment. Images are often also "standardised" in some way to aid later interpretation and generalisation of the statistical results, particularly important for group analyses.

4.4.5.1 Image reconstruction

Images were reconstructed from the raw time data and corrected for any ghosting artifact present. This can be a particular problem for EPI data and is seen as a lower intensity "shadow" of the real image that can cause artifacts to appear in the data where overlapping

occurs. From this point onwards all pre-processing and analysis steps were applied using the Statistical Parametric Mapping software package SPM2, which is specifically designed for the analysis of functional imaging time series data (http://www.fil.ion.ucl.ac.uk/spm/). The data from all three participants was modified so that the origin of the images, i.e. the point where the x, y and z co-ordinates are all equal to zero, and their orientation closely approximated that of the SPM template brain images. This helps to reduce error in the later pre-processing stage of normalisation, in which subject images can be warped to match a template that represents a standard brain.

4.4.5.2 Image realignment

If a participant moves their head during the scan session, as is likely, it can mean that data attributed to a single voxel may have actually been contributed by different voxels depending on the position of the brain within the scanner when each volume was acquired. Chapter 2 detailed the problem head movement poses for functional imaging experiments and the particular relevance to studies requiring overt verbal responses. To reduce the potential for motion related artifact contaminating the results motion correction was applied to the data using a rigid body transformation to rotate and translate volumes around their axis. Firstly, the values were calculated for the six parameters that describe the difference in position between images. Then the images were re-sampled to spatially align them with the first image volume acquired in the time series.

4.4.5.3 Slice timing correction

The fMRI scanner captures images of the brain in a series of sequential slices covering the required field of view. Consequently, data from different voxels across the brain is collected at slightly differing time points relative to the start of the volume of slices. Statistical analysis in SPM assumes that all slices within a volume were acquired simultaneously. Shifting the time series for each voxel to align it to the volume start time can effectively simulate this. Images were therefore corrected for the delay in acquisition between slices to ensure greater accuracy in the modelling of the HRF time course across slices.

4.4.5.4 Normalisation

The data was spatially normalised to minimise morphological differences between individual brains and allow for valid comparison of results between subjects and with results from previous studies. Each individual subject's brains were warped to match a template brain as closely as possible in order to place all subject data in the same standard 3D space.

4.4.5.5 Spatial and temporal smoothing

Images were spatially smoothed using a gaussian kernel with its full width at half maximum (FWHM) set to 8mm. The FWHM value sets the range of the blurring effect within the local neighbourhood. The signal within the gaussian kernel is effectively averaged across pixels, which can reduce the impact of irregular noise without removing any signal of interest and thus increasing the SNR. Smoothing also has the function of reducing any small structural differences between participants that survived the normalisation process.

The last process applied to the data before statistics were calculated was a high-pass temporal filter with the cut off set to 128s. This correction works on each voxel's time series to remove uninteresting low frequency variability without affecting the signal change associated with the time course of the experimental stimulations. Removal of the low frequency noise should improve the fit of the statistical model to the data and increase the liklihood of detecting significant activations.

4.4.5.6 Statistical analysis

The statistical approach used to interrogate fMRI data in this piece of research was a model driven one, where the expected pattern of haemodynamic response to the experimental paradigm is described and compared to the data acquired. Generally speaking, the better the fit between expected and actual changes in signal then the greater the probability that the activations are a true result of the experimental stimulation. Where there is more than one stimulus condition additional stimulus timecourses can be added to the model to be fitted to the data. The model also needs to include some estimation of error to account for the variability that can randomly occur in the haemodynamic response to the experimental conditions even in exact replications of the study with the same subject.

The timings of each stimulus application were inputted into the statistical model to describe the pattern that the functional imaging data is expected to contain. The resulting waveform was convolved with the standard canonical HRF so that the model looks more similar to the responses produced in the brain, resulting in a better fit to the hypothesised stimulus driven activation. No temporal or spatial derivatives were added to reduce the likelihood of including changes in signal caused by subject motion in the model (see Chapter 2 for a full explanation). The parameters that had been calculated to describe subjects' head movements in the motion correction stage were added to the model as regressors of no interest. This was done in order to remove any remaining variance from the model associated with these estimated movements. It is understood that if the task paradigm is correlated with movement in the scan cycle then there is a danger that using these parameters as regressors can lead to the removal of true signal from the statistical analysis. It was decided, however, that it is preferable to risk removing some true signal change than leave the statistical results possibly contaminated with motion related error and false positives.

Activation maps were generated for the main effect of the naming condition, i.e. a one-sided t test. The t statistics for the results analysis were generated using a significance level threshold of $p < 0.05$. Due to the vast quantity of statistical tests being performed across all voxels within the field of view a correction for multiple comparisons needed to be applied. Without performing such a correction it is likely that a large number of voxels will show activation purely by chance, leading to a high proportion of false positives in the analysis. For example, thresholding a set of 100,000 voxels where no true activation is present using an uncorrected p value of 0.05 would be expected to result in 5,000 false positives. In this pilot study the problem of employing multiple comparisons was controlled for using the family-wise error rate (FWE), i.e. the probability of any false positives occurring in the data.

The cluster size of the activations was limited to a mimimum of two voxels to prevent inclusion of random spikes in the data.

4.4.6 Results

Subjects achieved a mean accuracy rate of 97% for the naming task.

The motion plots generated from the estimated head motion in each session show small movements when each response was given, away from and back towards the original position of the head (see figure 5). These movements are not large, however, and are in accordance with findings from Barch et al (1998) and Huang et al (2001), that subjects producing overt verbal responses do not show significantly greater movement away from their initial head position than when producing covert responses.

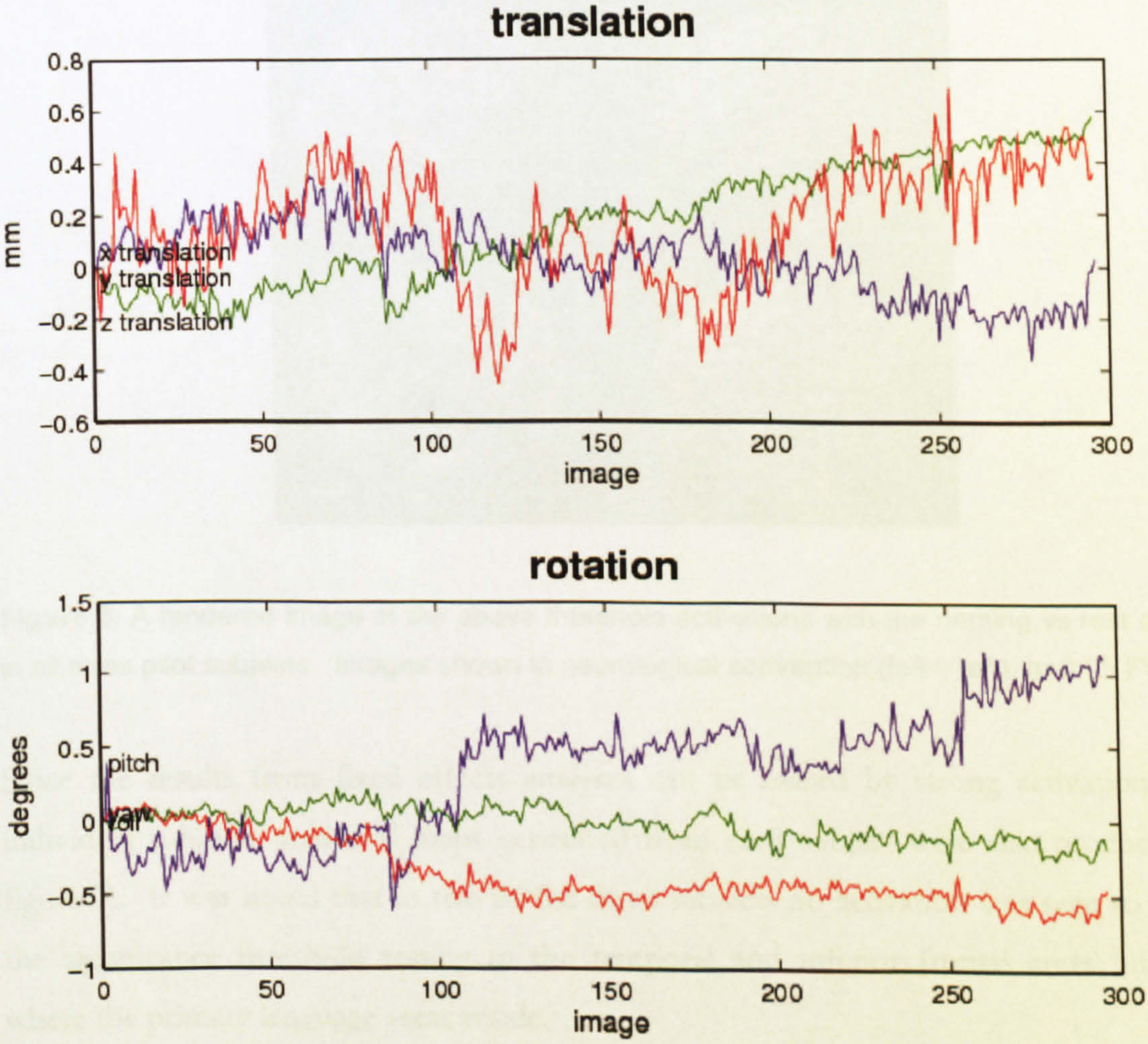

Figure 5. Sample plots of motion correction parameters, taken from the pilot study. Plots show the three transformations and three rotations applied to images during the realignment process.

Statistical maps for the contrast of naming versus rest were generated for the group of three subjects, using a fixed effects analysis. Figure 6 shows all areas with above threshold

values displayed on a standard 3D rendered brain. Areas showing significant activity levels include the visual cortex in the occipital lobe, superior temporal gyrus primarily in the right hemisphere and the motor cortex.

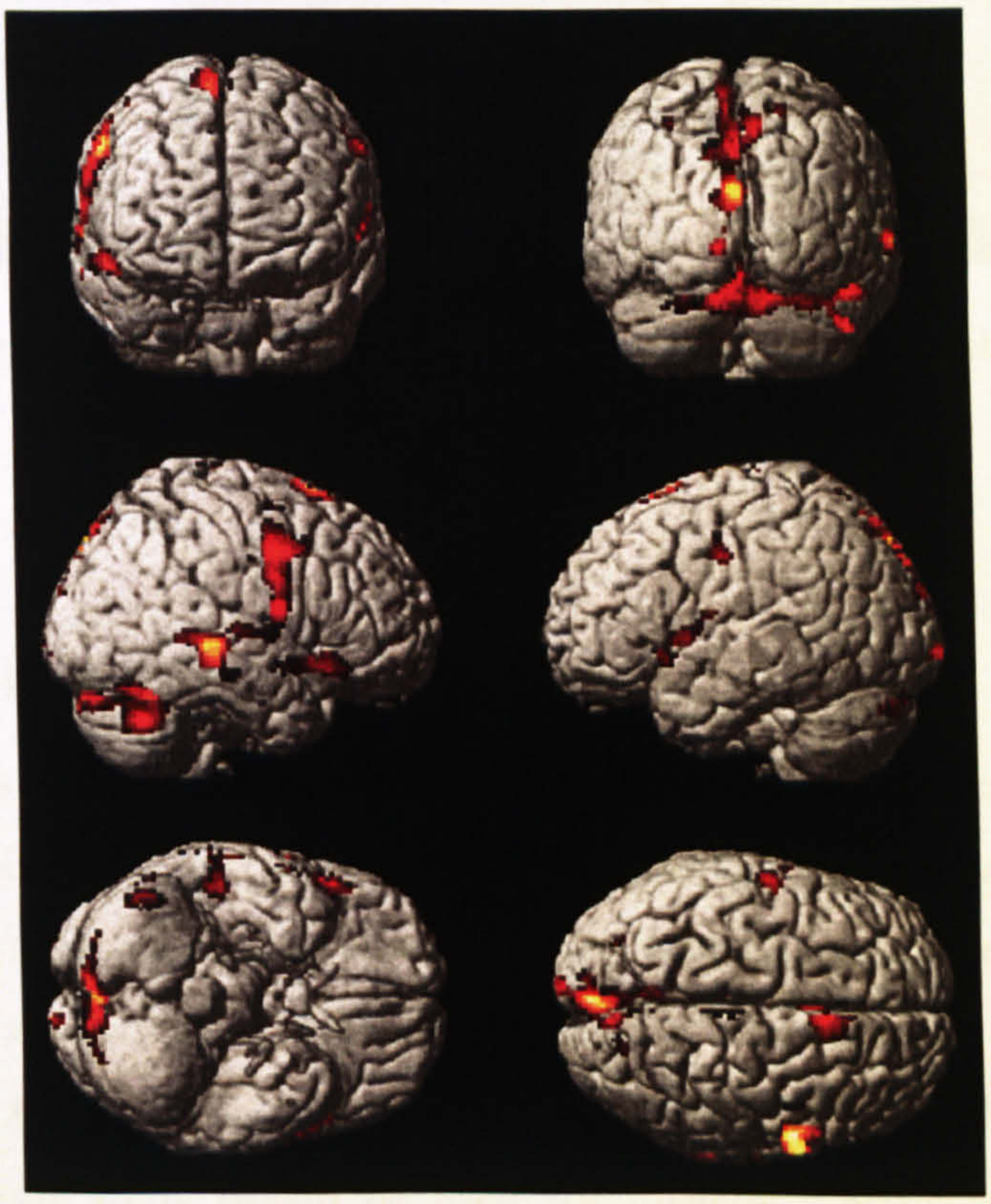

Figure 6. A rendered image of the above threshold activations with the naming vs rest contrast in all three pilot subjects. Images shown in neurological convention (left = left), p>0.05 FWE.

Since the results from fixed effects analyses can be biased by strong activations from individual subjects, statistical maps generated from each subject were also produced (see figure 7). It was noted that in two of the three subjects no activation was seen to survive the significance threshold testing in the temporal and inferior frontal areas bilaterally, where the primary language areas reside.

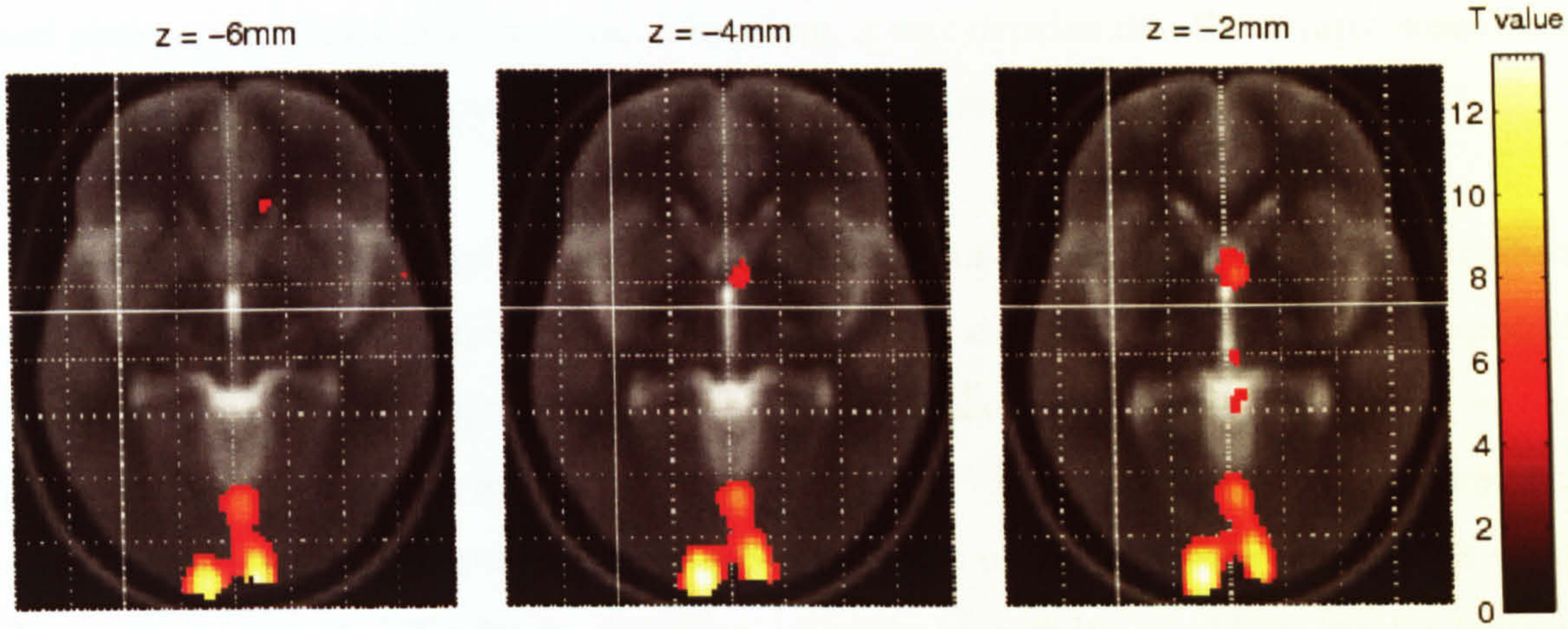

Figure 7. Slices from an individual subject showing activations and their intensities. P>0.05 FWE (left=left).

4.4.7 What went wrong?

Although it appears that some activations seen in the results are within the brain regions predicted by previous evidence into the neural correlates of naming they are not extensive and many expected areas are completely missing from the maps. Those areas may not have reached the significance threshold, which indicates a lack of power in this design to detect areas of neural activity. The results from this pilot are not substantial enough to recommend the same study design be used for imaging aphasic patients. There are, however, steps that can be taken to improve the situation.

The field of view imaged in this study covered the whole brain. In order to cover this area the thickness of the slices acquired and the time to repeat (TR) for a volume had to be relatively large for an EPI ER-fMRI paradigm. As a result the temporal and spatial resolution of the images are not as desirable as they could be, which may be a major factor in the weakness of the activations observed.

A simple way to increase the statistical power of an event-related design is to extend the number of task events within the experiment. This would provide more data points to be entered into the analysis and give the paradigm more sensitivity to detect signal changes. However, an important concern in this study is patient comfort and the restriction of excessive movement within the scanner. Adding a significant number of events that would make an impact on detection power would mean compromising with a patient's comfort

and ability to minimise head motion. Therefore, it was decided that this course would not be explored unless other potential solutions fail to be effective.

Three subjects is not a large amount to conduct a group analysis on. Additional subjects would provide more statistical power and improve the ability to detect activity in expected language areas when a significance threshold is applied. However, data from stroke patients needs to be analysed on an individual basis. Their structural and performance characteristics are very variable and so grouping their results together may not be valid. Also, the increased difficulty in detecting task-related activity in elderly stroke patients needs to be taken into consideration. For these reasons it is necessary for individual pilot results to demonstrate strongly significant patterns of activation that correlate with previous naming studies, so there can be confidence that the same experimental paradigm and scanning protocol can elicit statistically interpretable results in a more inhomogeneous population.

4.5 Pilot Study 2

4.5.1 Changes made

In light of the results from this pilot data an additional study was conducted, also using subjects with normal language ability. The task paradigm used was identical to that from the original pilot study, but with changes were made to the fMRI protocol in an attempt to boost the statistical power of the experimental design. The imaging field of view was narrowed from whole brain coverage to only acquiring data from the areas identified as being involved in the linguistic aspects of the naming process. This omitted the occipital lobe, posterior parietal and anterior frontal cortex (see figure 8).

As a result of the reduction in field of view the number of slices required for coverage could be reduced from 22 slices to 18. The time taken to acquire a volume of data was decreased accordingly from a relatively high 2.99s to 1.99s (TR=1.99s). Smaller voxel dimensions could also be chosen (3x3x6mm) to increase spatial resolution and significantly more volumes of data could be acquired within the time of the experiment. 454 volumes were entered into the analysis, increased from 296.

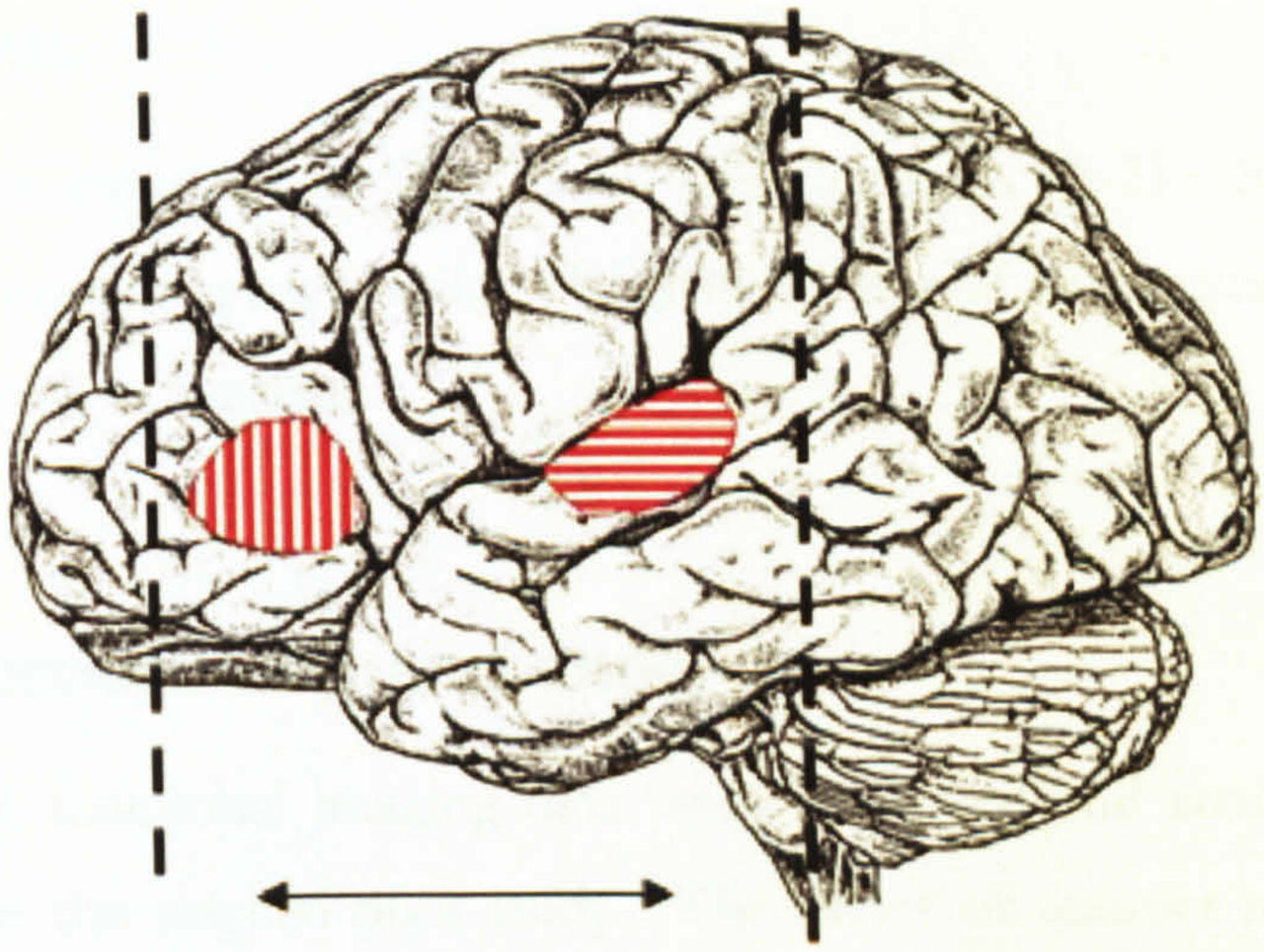

Figure 8. Field of view used in the second pilot study, with the positions of classical language areas, Broca's and Wernicke's areas, marked.

The aim of this second pilot study was to test whether alterations to the scanning protocol, described above, would improve the power to detect task related signal change in brain areas known to be involved in the object naming process. If this protocol is to be used to compare activation patterns arising from correct and error naming trials, one of the major aims of this research project, it needs to demonstrate a reasonable baseline level of statistical power for detecting signal change. This is so that when trials are separated into the two conditions (correct and error) the resulting decrease in power does not prevent major activations from reaching threshold levels of significance.

As for the first pilot study, it was predicted that functional imaging data would show increases in signal in areas involved in object naming, such as posterior inferior frontal cortex (including Broca's Area), anterior, posterior and inferior areas of the middle and superior temporal cortex (including Wernicke's Area and fusiform gyrus), inferior parietal cortex (including angular gyrus) and cingulate cortex. Obviously, the altered imaging protocol means that predicted activations in the primary visual cortex in the occipital lobe can no longer be acquired.

4.5.2 Subjects

Six subjects were recruited for this study, aged between 18-21. Subjects were all right-handed with normal language ability, English as their first language, and no history of neurological disease or damage.

4.5.3 Data processing and analysis

Behavioural and functional imaging data were processed and statistically analysed in the same way as for the original pilot study. The effect of scanner noise in the sound files containing subject responses was reduced using the scanner noise cancellation tool. Responses were then transcribed and checked against the stimulus presentation order. The fMRI images were reconstructed, realigned, slice timed, normalised, smoothed and temporally filtered during pre-processing according to the processes described earlier in the chapter.

Stimulus timings and subject head movement parameters were entered into the model to explain the data. Activation maps were again generated for the naming versus baseline contrast and thresholded at a significance level of $p > 0.05$, corrected for multiple comparisons using FWE.

4.5.4 Results

Areas activated by the picture naming condition included inferior frontal areas, superior temporal gyrus, fusiform gyrus, supramarginal and angular gyri and cingulate cortex. These areas were strongly activated bilaterally, but most prominent within the left hemisphere (see figure 9).

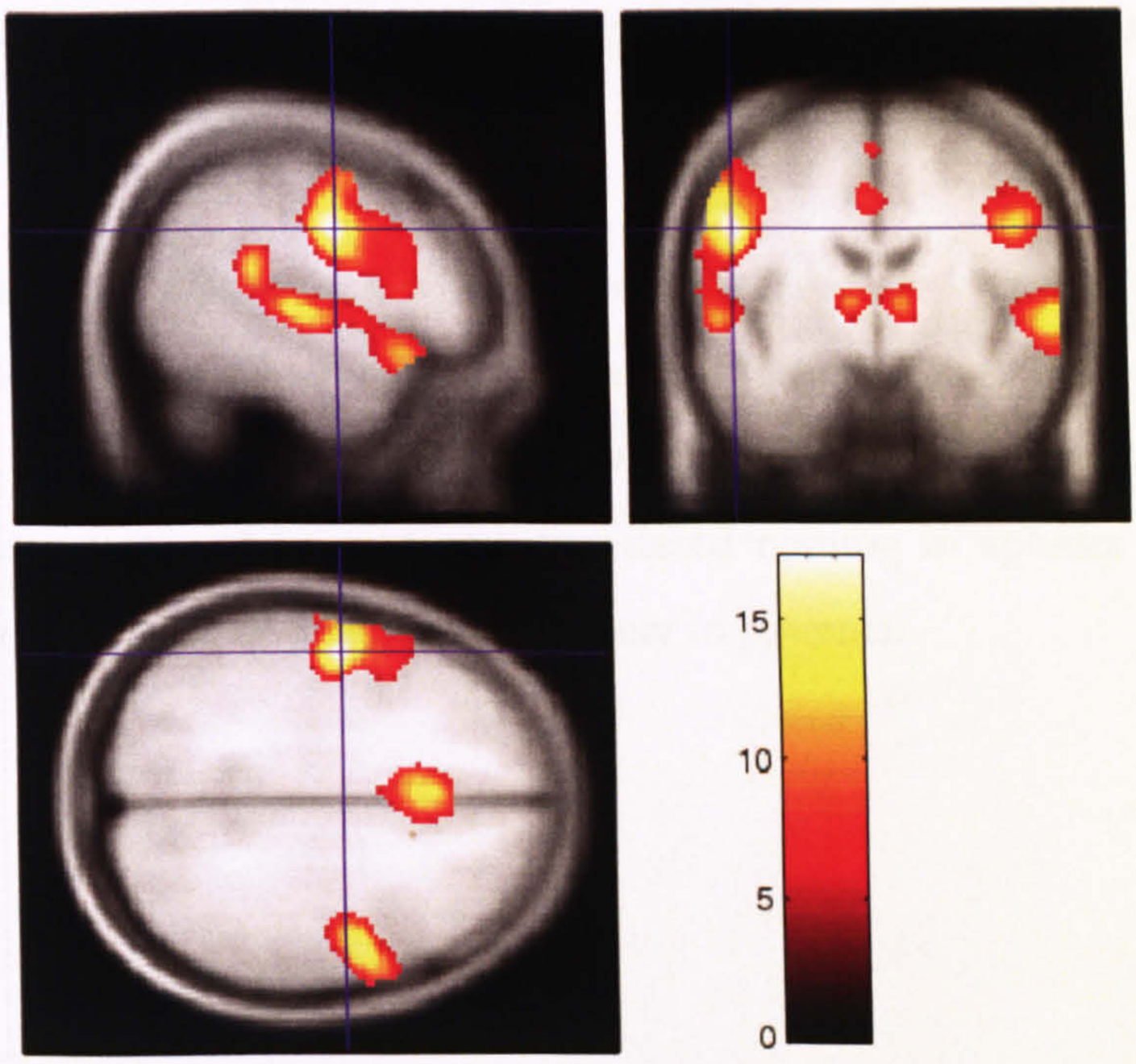

Figure 9. Activation map from pilot 2 group analysis overlaid onto the EPI template in SPM with signal intensity bar (p>0.05 FWE, left=left).

Individual analyses were also conducted for all 6 subjects. The activation maps for the group reflect individual results well, with similar areas reaching above threshold levels in all participants. The areas showing significant activation for naming pictures agree well with the results of published functional imaging studies of naming.

4.6 Conclusions

With the recognised difficulties in conducting an fMRI study of language abilities and disabilities in elderly stroke patients it is vital that the experimental task paradigm and scanning procedure be optimally designed. Although the initial pilot study was deemed to be not as efficient at detecting levels of neural activity as was necessary, the implementation of a new imaging protocol in the subsequent pilot produced much more successful statistical activation maps. The increase in spatial and temporal resolution and dataset size provided significantly more power to deliver results comparable with those from previous functional imaging studies of confrontation naming. It was judged that these results fulfilled the aims of the pilot work, by successfully trialling the fMRI protocol, experimental paradigm, equipment and participant experience to a satisfactory level. The hypothesis, that significant activations will be seen in inferior frontal cortex, areas of the

middle and superior temporal cortex, inferior parietal cortex and cingulate cortex, was also well supported by the statistical analysis of the data acquired in this second pilot study.

These new parameters will be used when scanning aphasic stroke patients to study what is occurring in their brains when they are producing correct and error responses to the picture naming task. It is hoped that this new study design will enable valid inferences to be made on the sources of normal and disordered naming in aphasia and how this can relate to the practical issue of promoting recovery in patients.

5 Patient Work

5.1 Introduction

The focus of this research is the experimental exploration of the production of errors in people with aphasia using fMRI. All other studies included within the scope of this project were completed either as part of the design process for the novel paradigm used in this fMRI experiment or to compliment and further inform the results achieved from this study.

The aims of this work were, firstly, to attempt to acquire usable event-related fMRI data from post-stroke patients with aphasia using continuous scanning with an overt language paradigm, and to compare correct and error trials for each subject. Prior to the start of this project, existing studies in the literature had not compared cortical activations associated with successful and unsuccessful naming in this way. The bulk of research concentrated on recovered language, or ignored error trials, which provides a blunt tool for investigating the relationship between task performance and brain activation patterns. Therefore, the second aim of this patient study was to tease out brain areas selectively involved in correct naming and those concerned with the production of naming errors.

It is hoped that results can inform current theories of how the brain recovers functional language after damage and provide a guide for the direction rehabilitative therapies should take. For example, if it is found that, as some studies suggest (Rosen et al, 2000; Karbe et al, 1998; Price and Crinion, 2005), the release of right hemisphere language homologues interferes with successful naming then this result would provide support for inhibitory treatments, such as the use of transcranial magnetic stimulation, to improve performance (Naeser et al, 2005). Behavioural rehabilitation techniques could also be used to encourage patients to use alternative task strategies that avoid the damaged areas. It is felt that this research could indicate a use for functional imaging technologies as part of normal diagnostic methods in speech and language therapy by potentially revealing where in the brain language processing is going awry.

Evidence in the current literature (reviewed in Chapter 1) led to the hypothesis that correct naming performance would be associated with processing within preserved language areas in the left hemisphere whereas the production of errors would coincide with an increase in activation in right hemisphere language homologues. If the reverse is true, such an observed increase in activation of language areas in the right hemisphere during correct naming would suggest that these areas are indeed providing functionally useful language processing.

5.2 Recruitment of Aphasic Volunteers

A major challenge for the functional imaging part of the research project was the recruitment of suitable volunteers to participate in both an initial behavioural screening session and a follow-on fMRI study. Interested volunteers were identified through referrals from county-wide speech and language therapy departments, to whom presentations were given regarding the research project, and through direct contact at a local aphasia support group (Aphasia Nottingham) where the project researcher worked as a volunteer. Several potential recruits were immediately excluded from consideration due to the presence of metallic implants that prevented them from being scanned safely. A further two volunteers were keen to help the research but were very reluctant to be scanned. In total 14 volunteers were recruited to partake in the research project and all were assessed within their home environment to ascertain whether they met the fMRI study requirements.

Due to the design of the study the exclusion criteria employed was necessarily strict. It was decided that participants should be at least six months post-stroke to allow their aphasic condition to stabilise prior to testing. The project was allocated limited time on the fMRI scanner and this meant that there may necessarily be a delay between home assessments and the follow-on scanning session. It was desirable to minimise the chance of any significant changes in the participant's aphasia during their involvement in the research. The aphasic participants were required to have an anomic difficulty as measured by the Graded Naming Test (McKenna & Warrington, 1983) and a shortened version of the fMRI activation task, designed to identify any specific naming problems with particular types of stimuli. The object decision task from the Visual Object and Space Perception battery

(Warrington & James, 1991) was administered to ensure that an apparent aphasic naming difficulty was not instead a manifestation of a perceptual problem. The Token Test (Boller & Vignolo, 1966) and the Pyramids and Palm Trees assessment (Howard & Patterson, 1992) were also used to provide additional information on the nature of the volunteer's aphasia.

From the 14 recruited and assessed, 5 aphasic participants continued to the scanning phase of the study. The main excluding factor for the remaining participants was a ceiling level performance on the naming assessments. Since the focus of this study is a comparison of correct and error responses to a naming task it is obviously not useful if no, or very few, errors are made. It was a recognised issue for recruitment that those patients who were motivated to respond to the invitation to take part in the research were also likely to be relatively competent in their language abilities. Those with more severe aphasia tended to be unwilling to participate from the outset or drop out at an early stage. This was a problem that proved difficult to overcome throughout the course of the project.

5.3 Case Studies

Hereafter, volunteers in this study are referred to by the initials AP (Aphasic Participant) and a number that denotes the order they were recruited and tested in.

AP	Age	Sex	Time since stroke	Screening Test Object Naming	Graded Naming Test	Pyramids & Palm Trees	Token Test	Object Decision
AP1	58	M	13 mths	16/30	11/30	50/52	11/32	19/20
AP2	62	M	4 yrs 5 mths	15/30	12/30	42/52	22/32	20/20
AP3	67	F	3 yrs	18/30	14/30	51/52	15/32	19/20
AP4	37	M	2.5 yrs	29/60	1/30	39/52	0/32	18/20
AP5	81	M	10 yrs & 2 yrs	36/60	10/30	49/52	7/32	18/20

Table 1. Details of the demographics and behavioural assessment results of the five aphasic participants.

The participants consisted of 4 males and 1 female with an age range of 37-81 and a mean age of 61 ± 14 years (mean ± standard deviation). All had English as their first language and were right handed, as tested by the Edinburgh Handedness Inventory (Oldfield, 1971). Table 1 shows the results of the language and cognition assessments for all the participants. All had suffered a left hemisphere middle cerebral artery stroke resulting in their aphasia. Figures 1 and 2 show structural data for the 5 participants revealing the site and extent of their lesion. The images for AP3-5 are T1-weighted 3D Turbo Spin Echo (TSE) structural MRI data, however, due to the resources available at the time of their involvement in the research such high quality data was not able to be obtained for AP1 and AP2. Attempts were made to recall these two patients for detailed structural images to be acquired, however, AP1 could not be contacted and AP2 had sadly died. Their lesions therefore are shown on whole head EPI data collected during the original scan session.

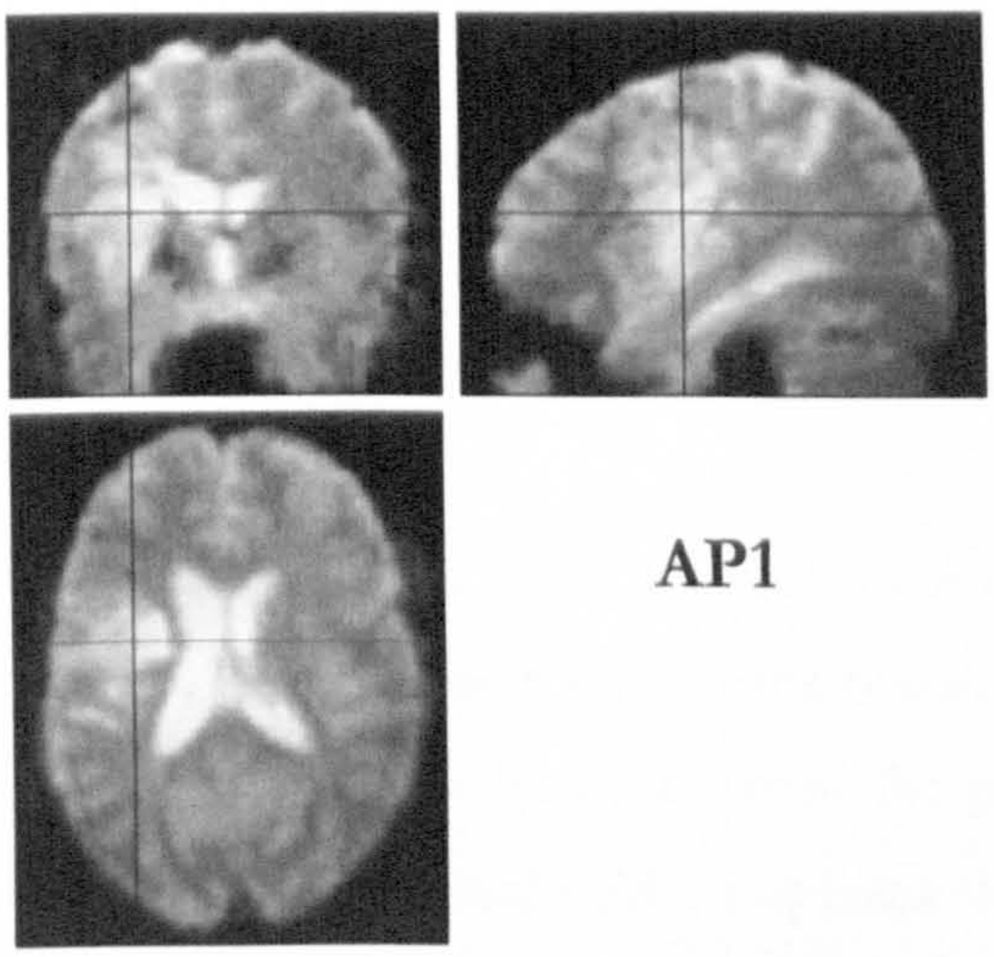

Figure 1. Whole head EPI images of AP1 (left) and AP2 (right) showing the site of their stroke lesion.

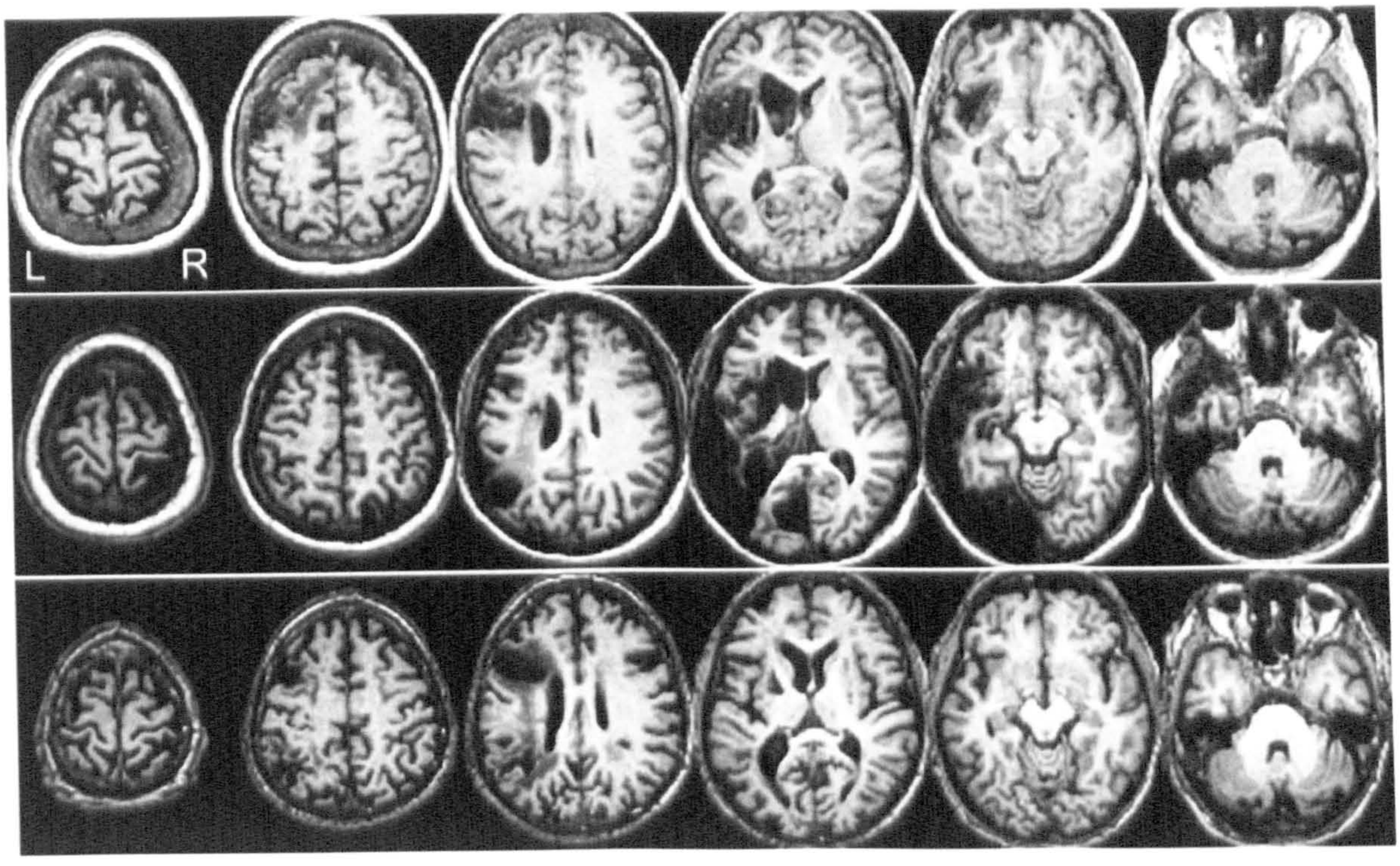

Figure 2. T1-weighted 3D TSE Structural images showing the stroke damage for AP3 (top), AP4 (middle) and AP5 (bottom).

5.3.1 AP1

AP1 was 58 at the time of recruitment into the study. He had suffered a large left middle cerebral artery infarct 13 months before being behaviourally assessed and scanned. When scanned the stroke lesion covered the precentral gyrus and posterior inferior frontal gyrus (see Figure 1). Initially AP1's aphasia was very severe with phonological output problems, but he had recovered the ability to communicate with non-fluent, slow and effortful speech. Significant word finding difficulties in every day conversation and in confrontational naming tests remain and speech is littered with long pauses whilst searching for the correct word. The scores achieved on the naming assessment tests revealed his word finding difficulties (see Table 1), but good performance on the Pyramids and Palm Trees test show that semantic access was intact. Errors on the naming tests tended to be in the form of omissions or descriptions of the target name. Some phonological errors were also made, for example, a picture of a tutu was named "foo-foo". A good score on the Object Decision Test indicated intact visual object recognition with performance on the Token Test revealing a difficulty in following more complex verbal commands.

5.3.2 AP2

AP2 suffered from a left hemispheric ischaemic stroke 4 years and 5 months prior to inclusion in the study leaving him with damage in the areas of inferior frontal gyrus, superior temporal gyrus and inferior precentral gyrus (see Figure 1). He was 62 when recruited. Initial stroke-related deficits included dysphagia, right hemiplegia and severe aphasia. AP2 had no verbal output following his stroke, with comprehension limited to slow, simple speech. During the course of recovery speech developed into long utterances with stilted phrasing and intonation. Errors on standard assessments used in speech therapy showed a predominance of phonological type errors, for example, trousers named as "twouzez". Behavioural assessments conducted as part of this study showed AP2 to have unimpaired visual recognition of objects, ability to follow simple commands easily, some semantic impairment and a significant deficit in confrontational naming.

5.3.3 AP3

AP3 displayed fluent speech with frequent use of jargon at point of contact, 3 years post-stroke. She performed near perfect on the Pyramids and Palm Trees test, showing no deficit in semantics, and had good comprehension of simple spoken commands. Performance on the Graded Naming Test and the screening object naming test indicated an object naming impairment of a suitable level for inclusion in the study. Structural images of AP3's stroke lesion showed damage to Broca's area and premotor areas in the prefrontal cortex (see Figure 2).

5.3.4 AP4

AP4 suffered an ischaemic stroke as a result of a left carotid artery dissection 2.5 years before participating in the study. Structural images showed a large lesion covering much of Wernicke's area and occipital and temporal lobes (see Figure 2). This participant was the youngest aphasic patient tested, at 37 years. Since his stroke his speech had developed to the point where he could produce fluent sentences, though with some word finding difficulties in conversation and sometimes showing a lack of awareness of errorful utterances and would neglect to correct them. The assessment scores in Table 1 show an impairment in semantic access from pictures and a difficulty in understanding simple oral commands. Performance on the Graded Naming Test was near floor level, however, AP4

managed to correctly name half of the picture stimuli used in the screening object naming test.

5.3.5 AP5

AP5 had suffered two strokes prior to scanning. His first occurred 10 years before contact and his second 2 years before participating in the study. Damage included Broca's area and premotor areas (see Figure 2). His speech was very effortful and non-fluent with word finding difficulties in everyday conversation. Naming performance was shown to be impaired to a suitable level for inclusion in the study, with a good score on the Pyramids and Palm Trees Test implying that semantics were not the root of his naming problem.

5.4 Methods

5.4.1 Naming paradigm

The activation task used was identical to that developed during the pilot work. Sixty black and white line drawings from the Snodgrass and Vanderwart (1980) stimulus set were presented to the participants at the rate of one image per fifteen second cycle. The specific set of stimuli used was individually tailored for each subject according to their anomic difficulties, as observed in the screening assessment tests. The images were viewed for one second, with a fourteen second fixation cross between presentations. The total runtime of the paradigm was fifteen minutes. Responses were collected using the scanner safe microphone setup described in Chapter 4.

The paradigm was managed and presented using Presentation (http://www.neurobs.com). Digital recording of the speech output was also controlled within this software. AP1, AP2 and AP3 wore prism glasses which redirected their vision outside the scanner bore to a screen on which the stimuli were shown. For AP4 and AP5 the images were delivered via Avotec "Silent Vision" eye pieces (http://www.avotec.org) mounted onto the head coil.

5.4.2 fMRI protocol

fMRI data for AP1, AP2 and AP3 was collected using a 3T scanner custom built for the Sir Peter Mansfield Magnetic Resonance Centre at the University of Nottingham. During the

project this scanner was replaced by a Philips 3T Achieva magnet and imaging of AP4 and AP5 was performed on this system using an 8-channel SENSE head coil. Prior to the collection of this data some pilot work was carried out using the new Philips scanner with healthy, unimpaired volunteers to ensure that the naming paradigm and the developed scanning protocol would still provide usable data. Acquisition of EPI image volumes of the whole head was trialled to determine if gaining extra cortical coverage was viable within the scope of the paradigm. However, tests showed that the scanner acoustics associated with the whole head scanning protocol interfered with the speech recording to an unacceptable level. It was also thought beneficial to use parameters as similar as possible to those used with earlier participants to make the overall project results easier to interpret.

For all participants T2* weighted EPI images were acquired using 18 coronal slices covering temporal, anterior parietal and posterior frontal cortex (see figure 8 in Chapter 4 for a diagram showing the field of view) at a resolution of 3 x 3mm in-plane, 6mm slice thickness, TE = 35ms and a sense factor of 2 for AP4 and AP5. 454 volumes were collected with a repetition time (TR) of 1.99s for AP1-3 and 2.00s for AP4-5. An EPI volume of the whole head was also collected using the same resolution as the fMRI images.

In addition to the functional data, T1 weighted anatomical images were acquired for AP3-5. As previously stated, detailed structural images were not able to be acquired for AP1 and AP2. Using the later Philips magnet sagittal 3D TSE images of AP3, AP4 and AP5 were collected in 4.5 minutes with 1mm isotropic voxels and a sense factor of 2.

5.4.3 Data processing and analysis

5.4.3.1 Spoken responses

To promote clarity of the recorded responses a sound cancellation tool (Cusack et al, 2005) was used post-hoc to remove excess scanner noise from the sound file before responses were transcribed and categorised. The cancellation tool takes advantage of the regularity of the noise generated by the scanner. It estimates the TR of the scan sequence and uses this to extract a single cycle. The onsets of each additional cycle in the sound file are identified and the noise generated by the scanner in these cycles is estimated. The output is then simply subtracted from the complete sound data leaving the voice recording intact.

Participants' verbal responses were transcribed and identified as either a correct or error naming response. In line with previous studies of confrontation naming in aphasics (Dell et al, 1997; Caramazza et al, 2000; Capitani & Laiacona, 2004) a response was coded as correct if it matched the target word or its synonyms. All other responses were classed as errors and, for the purpose of behavioural analysis only, were further separated into the following categories:

Semantic – The response is related to the target name by meaning.

Formal – The response shares some phonemes with the target name.

Mixed – An error that fulfils the criteria for both semantic and formal categories.

Unrelated – No relationship exists between the response and the target.

Non-word – The given response cannot be found in the vocabulary.

No response – No response is offered.

Other – Consisting of multiple word or part responses and descriptions.

Additional categories of semantic/visual and unrelated/visual were introduced to the above error classification system to account for those responses where, while they fitted the definitions for semantic or unrelated errors, it could not be discounted that the errors had arisen due to the visual similarity between the target and the given response.

5.4.3.2 fMRI data

Pre-processing and statistical analysis of the data from each aphasic participant was completed using SPM5 (http://www.fil.ion.ucl.ac.uk/spm/).

fMRI images were slice timed and realigned as in the pilot studies (see Chapter 4 for descriptions of the pre-processing stages). Due to the patient data being considered on an individual basis and the desire to locate activations with reference to each patient's stroke lesion the additional pre-processing stage of co-registration was introduced to the workflow. Co-registration was used to align different brain images to allow for the results from statistical analysis of the functional images to be overlaid onto higher resolution or more complete images. Data for AP1-2 were co-registered to the whole head volume acquired. Data for AP3-5 was co-registered to their anatomicals and then normalised to MNI space. All patient data was spatially smoothed using a gaussian kernel with a 5mm

FWHM and finally a high-pass temporal filter was applied with a 30s cut-off to remove low frequency variability.

The processed data was modelled using a general linear model (GLM) design matrix with events separated into the two conditions of correct and error based on the participant's recorded naming performance. The data was modelled with a 1s on period that started at the exact speech response times and convolved with the canonical haemodynamic response function (HRF). The individual speech response times were included in the model to account for the variability of speech onsets in the aphasic participants' task performance. Temporal derivatives were added to the design matrix to account for any delay in haemodynamic response relative to the speech response.

Participants' data was analysed on an individual basis due to the variability of their stroke lesions. Statistical parametric maps (SPMs) of activated brain areas were generated for the two conditions of correct and error naming responses ($p < 0.05$, using family-wise error (FWE) correction for multiple comparisons). To reduce the problem of performing multiple comparisons across all voxels (see Chapter 4 for an explanation of the multiple comparisons problem) a small volume correction was applied to the data. This method limits the number of voxels that are considered for the statistical analysis by redefining the search volume to encompass only the brain areas of interest. This was done by creating a statistical map within SPM5 of all activation elicited in the experiment, representing a combination of the correct and error conditions, loosely thresholded at $p < 0.001$ uncorrected. The resulting image was saved as a masking image where activation is shown as white and everything else is made black. This mask was then applied to the contrasts of conditions to exclude areas of no interest from statistical consideration, reduce the number of tests performed on the data therefore decreasing the number of false positives likely to be present in the resulting activation map.

The contrasts (error > correct) and (correct > error) were formed ($p < 0.001$, uncorrected, masked by an inclusive mask of (error and correct) conditions at $p < 0.001$, uncorrected) to show areas selectively activated when each type of naming response was made.

5.5 Results

5.5.1 Naming performance

Patients produced an overall mean of 36 (SD = 5.50) correct naming responses and 23 (SD = 5.76) errors. Descriptive statistics for the patients' correct and error naming response times are presented in Table 2. Timings show that patients were generally quicker when producing a correct name, though patients' response times in both correct and error trials tended vary across a large range.

	Correct Responses			Error Responses		
	Mean (SD)	Range	N	Mean (SD)	Range	N
AP1	2.43 (1.20)	1.39-7.62	37	3.68 (1.84)	1.98-7.23	23
AP2	1.30 (0.32)	0.71-1.95	34	1.50 (0.49)	0.79-2.76	25
AP3	1.42 (0.36)	0.95-3.05	44	1.82 (0.66)	1.18-3.39	15
AP4	3.39 (1.39)	1.12-7.67	29	4.95 (2.79)	1.01-12.53	31
AP5	3.12 (1.09)	1.09-4.58	38	2.94 (0.92)	1.45-4.54	22

Table 2. Speech onset times for correct and error naming responses measured in seconds and the number of corrects and errors produced by each patient out of sixty trials.

Figure 3 shows the mean response times of correct and error naming responses for each patient individually. AP2 was fastest to respond in both conditions, with reaction times closer to the normal scale. AP4 was the slowest with the greatest ranges of times. It is interesting to note that AP5 produced errors slightly faster than correct responses.

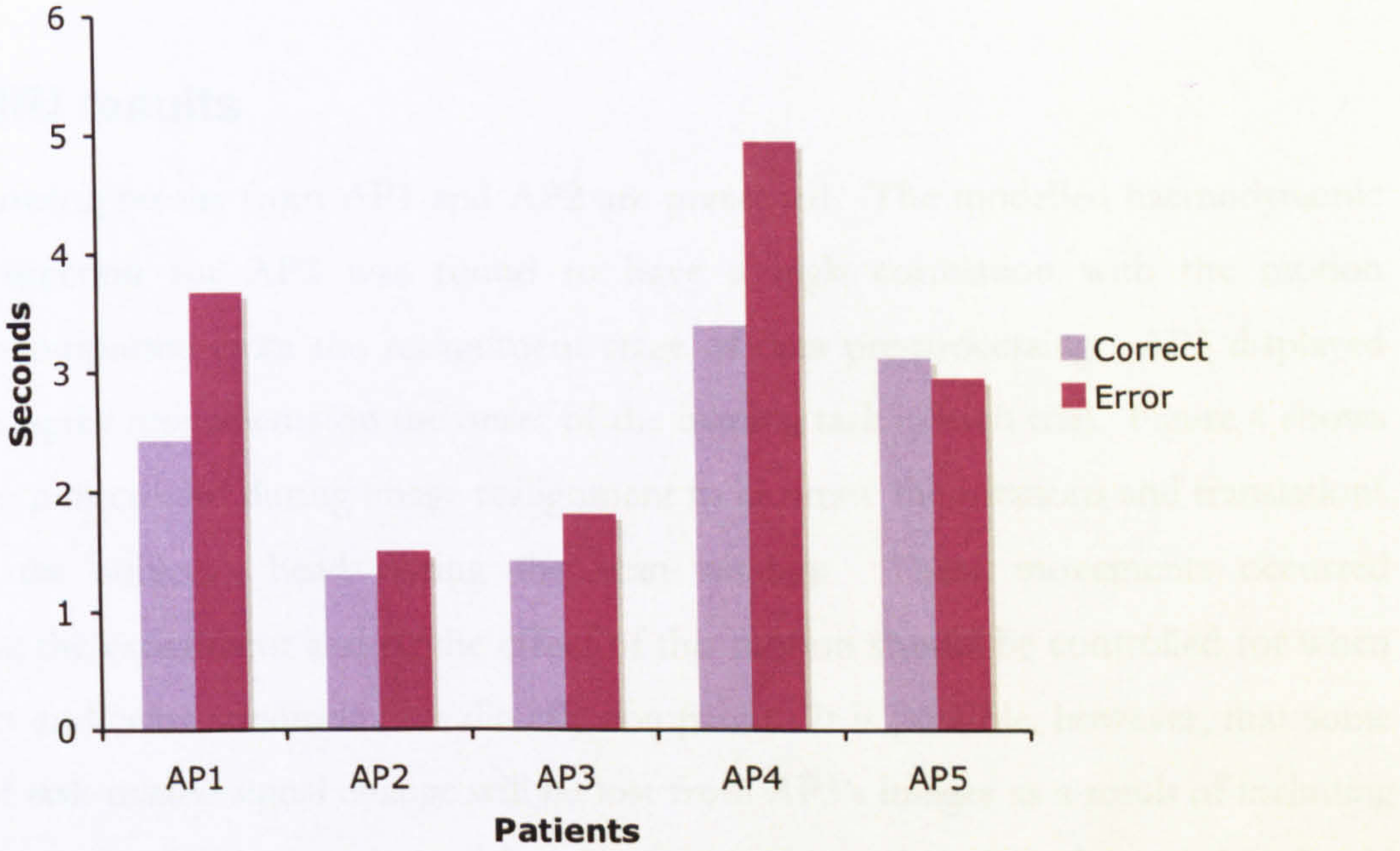

Figure 3. Mean speech onset times for correct and error naming responses across all five patients measured in seconds.

When error responses were categorised into different error types it was found that patients showed a variation in error patterns (see Table 3). AP1 tended to inhibit a response rather than produce an error, whereas AP2 predominantly produced errors of type "other", which included partial responses and descriptions of the target item. AP3 showed a spread across most of the error types and AP4 and AP5's errors were mostly semantic in nature.

Error Type	AP1	AP2	AP3	AP4	AP5	%
Semantic	3	4	4	18	8	32%
Semantic/Visual	3	0	2	0	0	4%
Formal	0	3	1	0	1	4%
Mixed	0	0	2	2	0	3%
Unrelated	0	0	1	1	0	2%
Unrelated/Visual	1	0	0	0	0	1%
Non-word	3	2	1	1	3	9%
No response	13	1	0	0	5	16%
Other	0	16	4	9	5	29%

Table 3. Breakdown of the number of different error types made by the aphasic participants, with the percentage of error trials represented by each error type.

5.5.2 fMRI results

Images showing results from AP1 and AP2 are presented. The modelled haemodynamic response function for AP3 was found to have a high correlation with the motion parameters outputted from the realignment stage of data pre-processing. AP3 displayed very sharp, spiky movements on the onset of the naming task in each trial. Figure 4 shows the motion plot created during image realignment to illustrate the rotations and translations made by the subject's head during the scan session. These movements occurred throughout the experiment and so the effect of this motion should be controlled for when the correct and error conditions are directly compared. It is possible, however, that some amount of task-related signal change will be lost from AP3's images as a result of including the motion parameters as regressors of no interest in the statistical model and this should be taken into consideration when interpreting the fMRI results. This problem did not affect AP4 and AP5 to the same extent and with the enhanced localisation of activated areas in relation to the stroke lesion that their detailed structural images affords, data analysis has concentrated on AP4 and AP5's results.

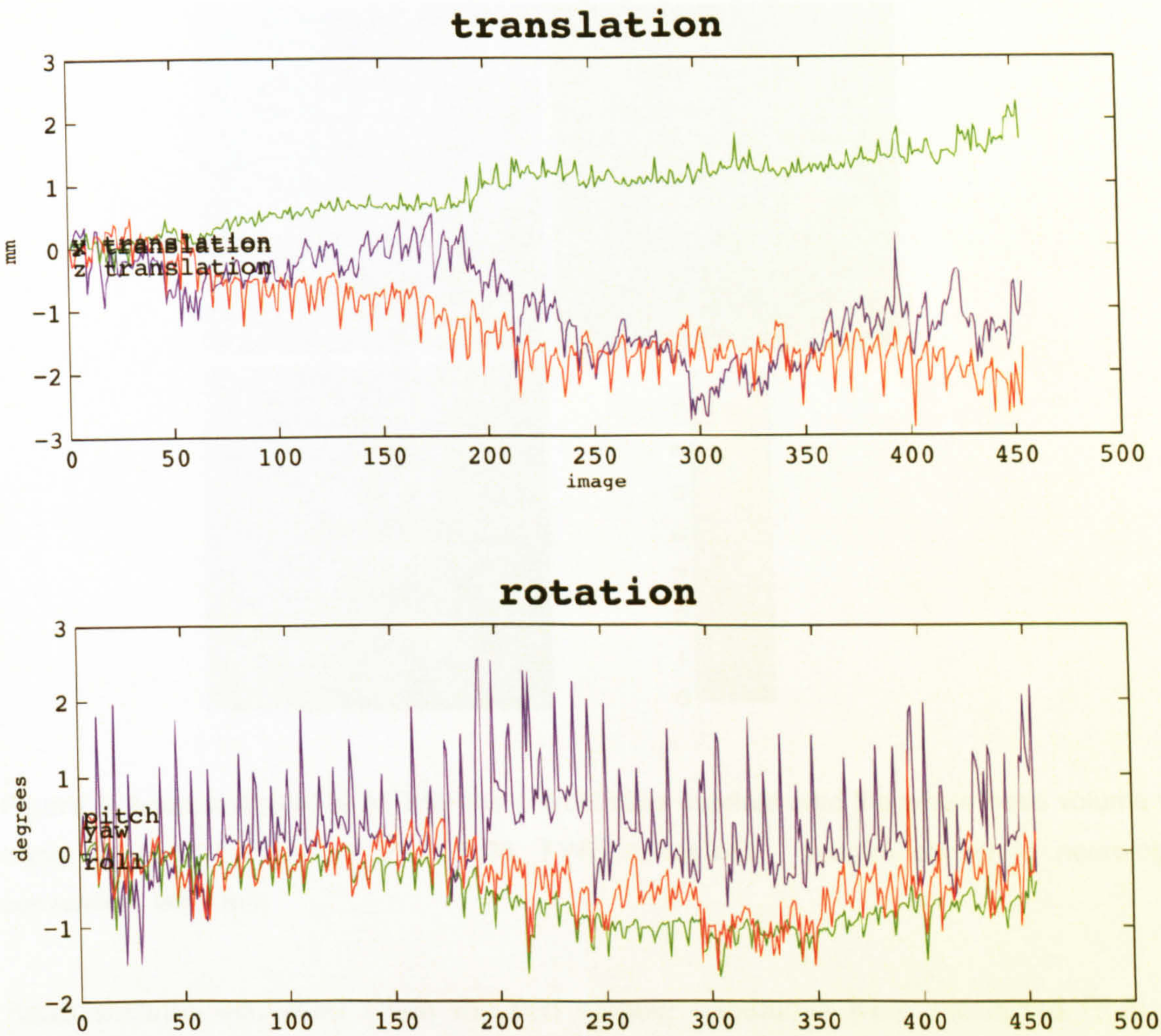

Figure 4. Motion plots for AP3 describing the estimated head motion in the x, y and z directions (translations in mm) and the roll, pitch and yaw rotations measured in degrees.

5.5.2.1 AP1

AP1's stroke lesion covered the left precentral gyrus and posterior inferior frontal gyrus. Few brain areas achieved significance for correct naming trials in AP1 (see Figure 5) with the global maxima including the right homologue of Broca's area (BA 44) in the right inferior frontal gyrus.

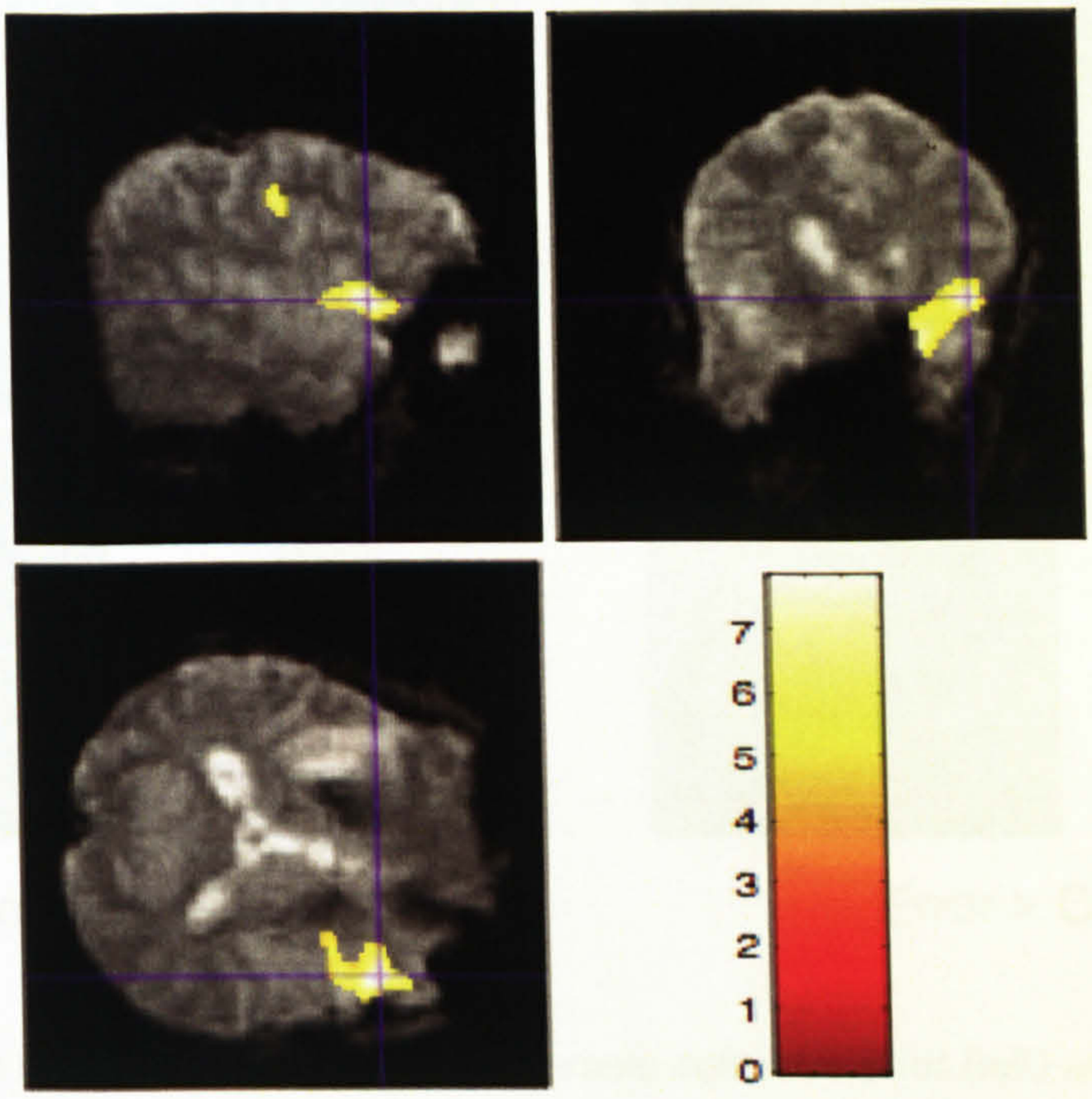

Figure 5. Areas activated in AP1 for the correct trials overlaid onto the whole head volume with signal intensity bar shown (P < 0.05, FWE corrected). Images shown in neurological convention, left = left.

Areas showing activation when the two naming conditions were subtracted from one another are shown in Figure 6. Few voxels achieved significance for the (error > correct) contrast (p > 0.001, uncorrected). One significant cluster of activation was shown to be selectively active in successful naming and is sited medial to the posterior portion of the right inferior frontal cortex, extending to the right superior temporal gyrus.

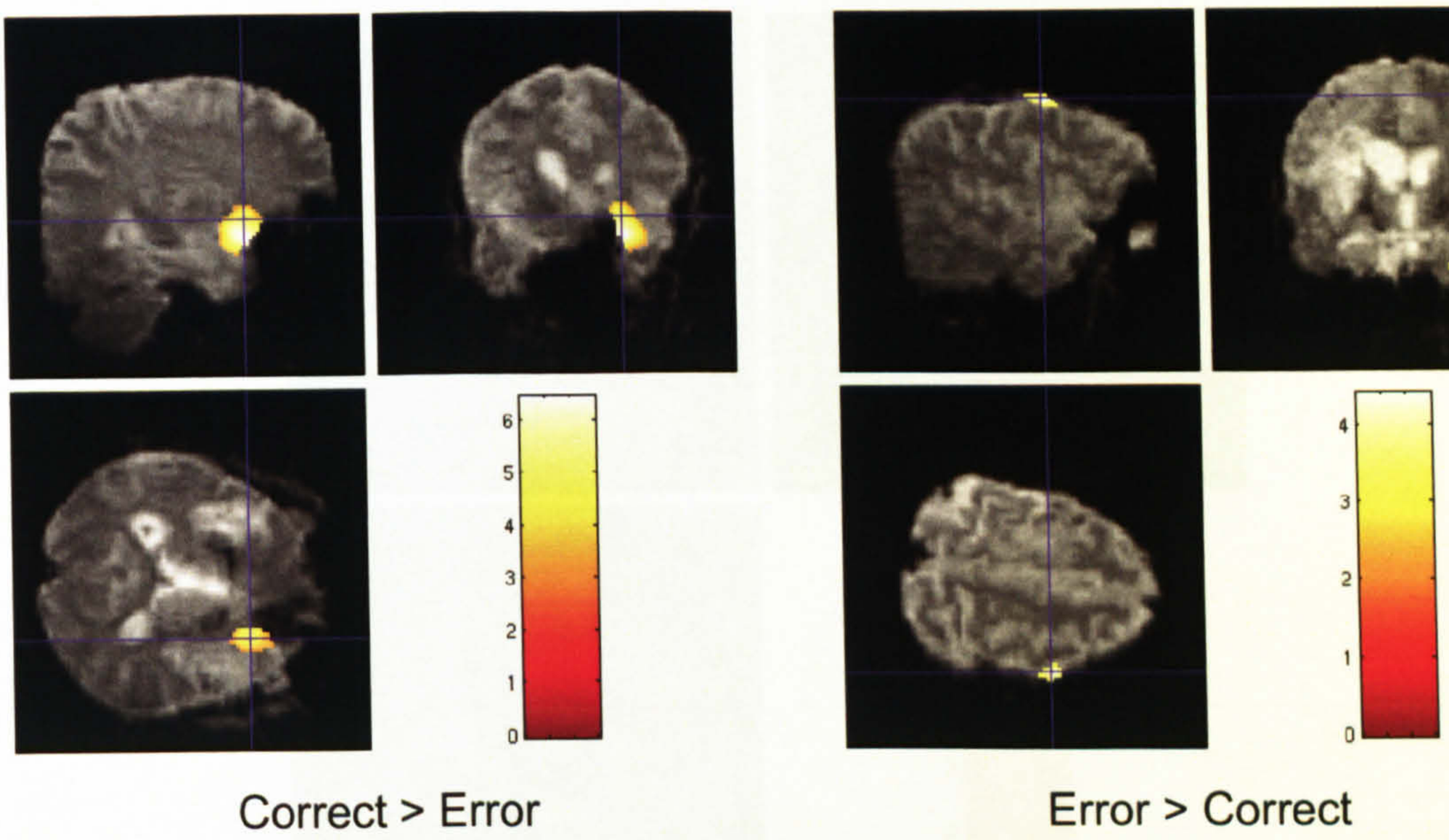

Figure 6. Areas activated in AP1 for the contrasts correct > error (left) and error > correct (right) overlaid onto the whole head volume with signal intensity bar shown (P < 0.001, uncorrected). Images shown in neurological convention, left = left.

5.5.2.2 AP2

Figure 7 shows areas active during correct picture naming for AP2, whose lesion included areas of the left inferior frontal and superior temporal cortex. Activation was seen in the right homologue of Wernicke's area in the right superior temporal gyrus

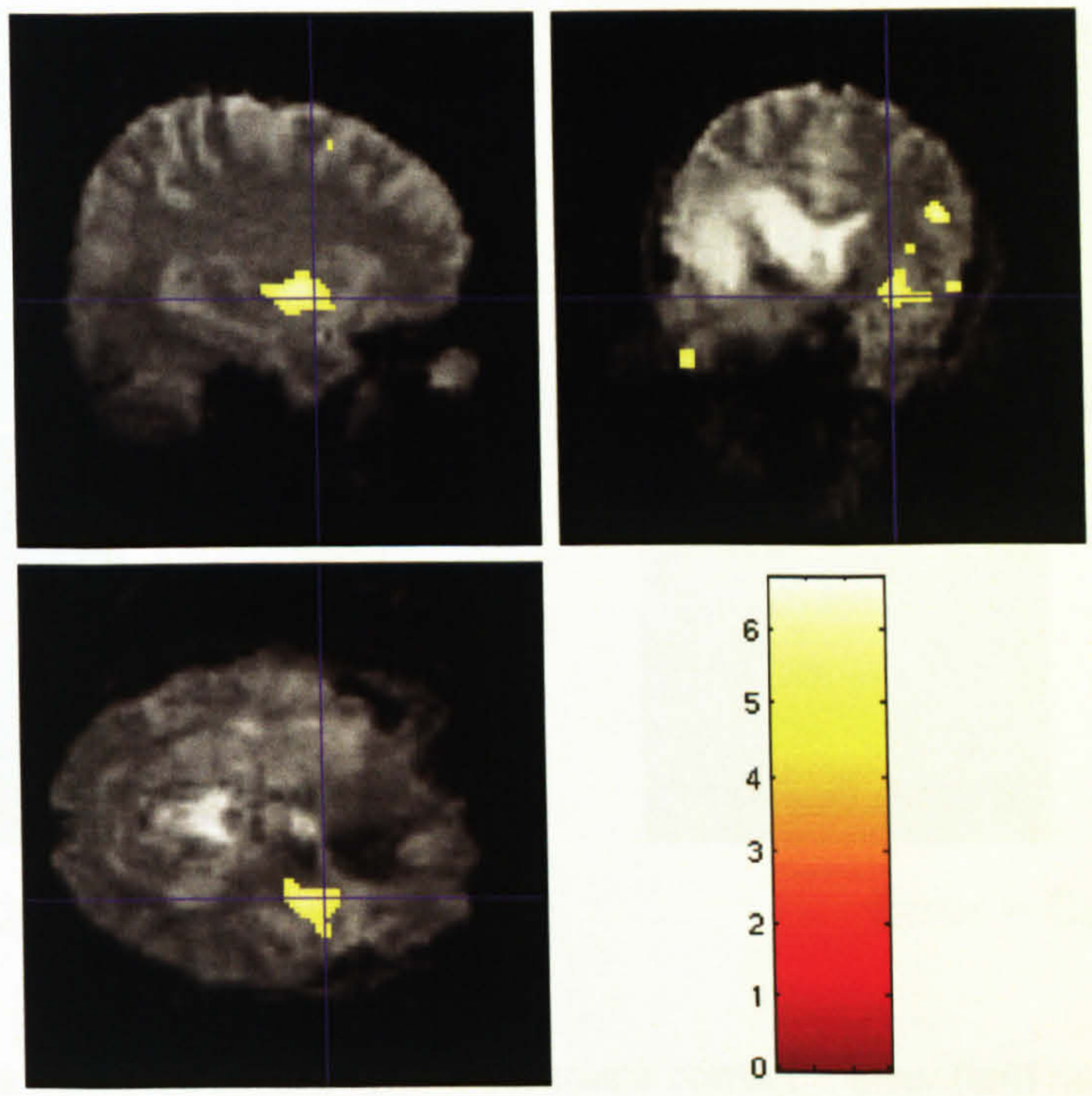

Figure 7. Areas activated in AP2 for the correct trials overlaid onto the whole head volume with signal intensity bar shown (P < 0.05, FWE corrected). Images shown in neurological convention, left = left.

Few areas showed activation when (correct > error) and (error > correct) contrasts were performed (see Figure 8). An area of the left postcentral gyrus showed greater significance for the error compared to correct trials. Other activations that were seen may be artifactual in nature.

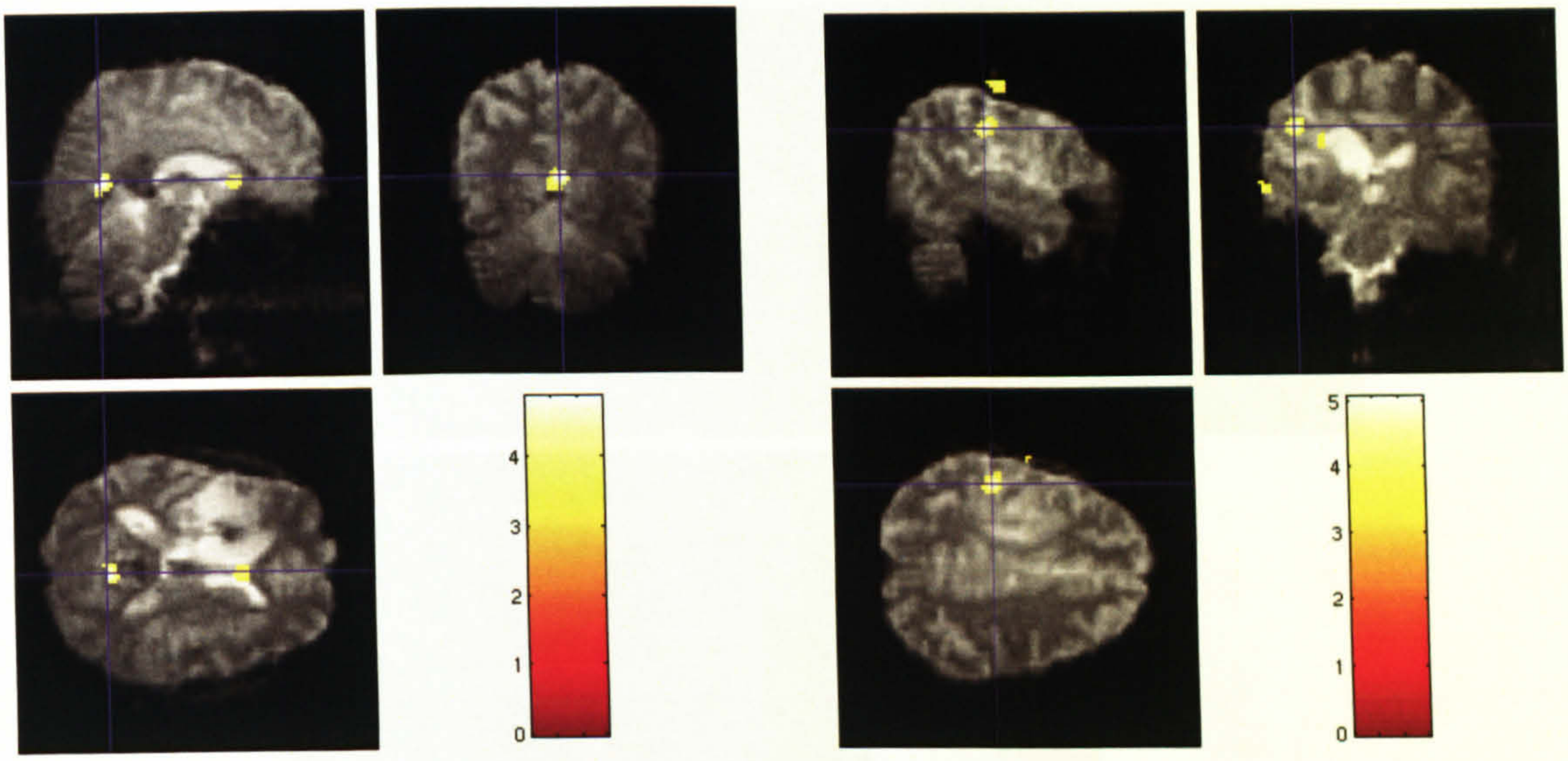

Figure 8. Areas activated in AP2 for the contrasts correct > error (left) and error > correct (right) overlaid onto the whole head volume with signal intensity bar shown ($P < 0.001$, uncorrected). Images shown in neurological convention, left = left.

5.5.2.3 AP3

Areas reaching significance for the correct condition in AP3 include the middle and superior temporal gyri in the left hemisphere, adjacent to the lesion site covering left Broca's and premotor areas (see Figure 9).

An area in the left middle temporal gyrus showed selective activation for the correct trials compared to the error trials ($p < 0.001$ uncorrected) and for the reverse contrast areas more significant for error responses were right Broca's area homologue, right premotor area and anterior cingulate gyrus (see Figure 10).

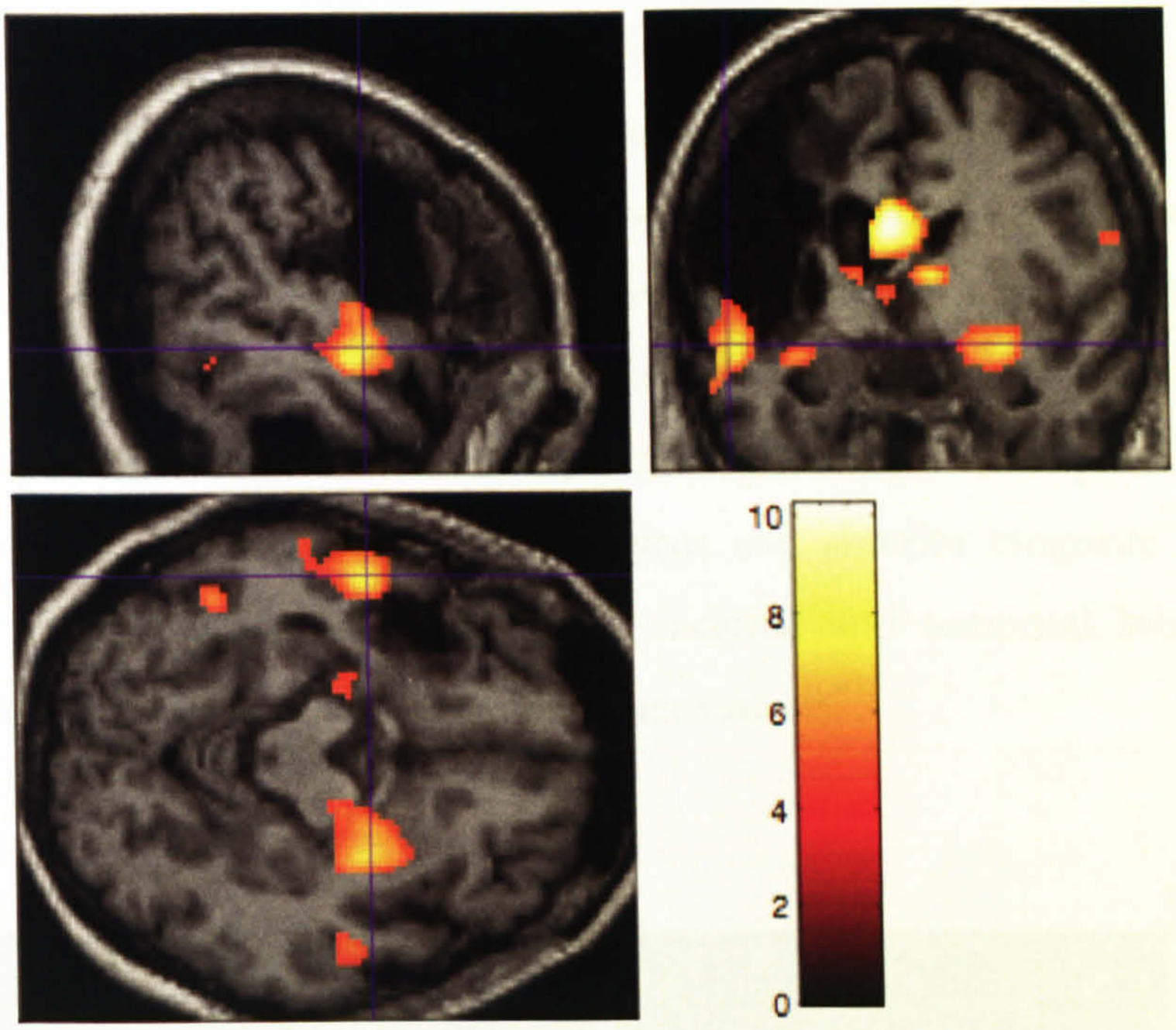

Figure 9. Areas activated in AP3 for the correct trials overlaid onto the structural image with signal intensity bar shown (P < 0.05, FWE corrected). Images shown in neurological convention, left = left.

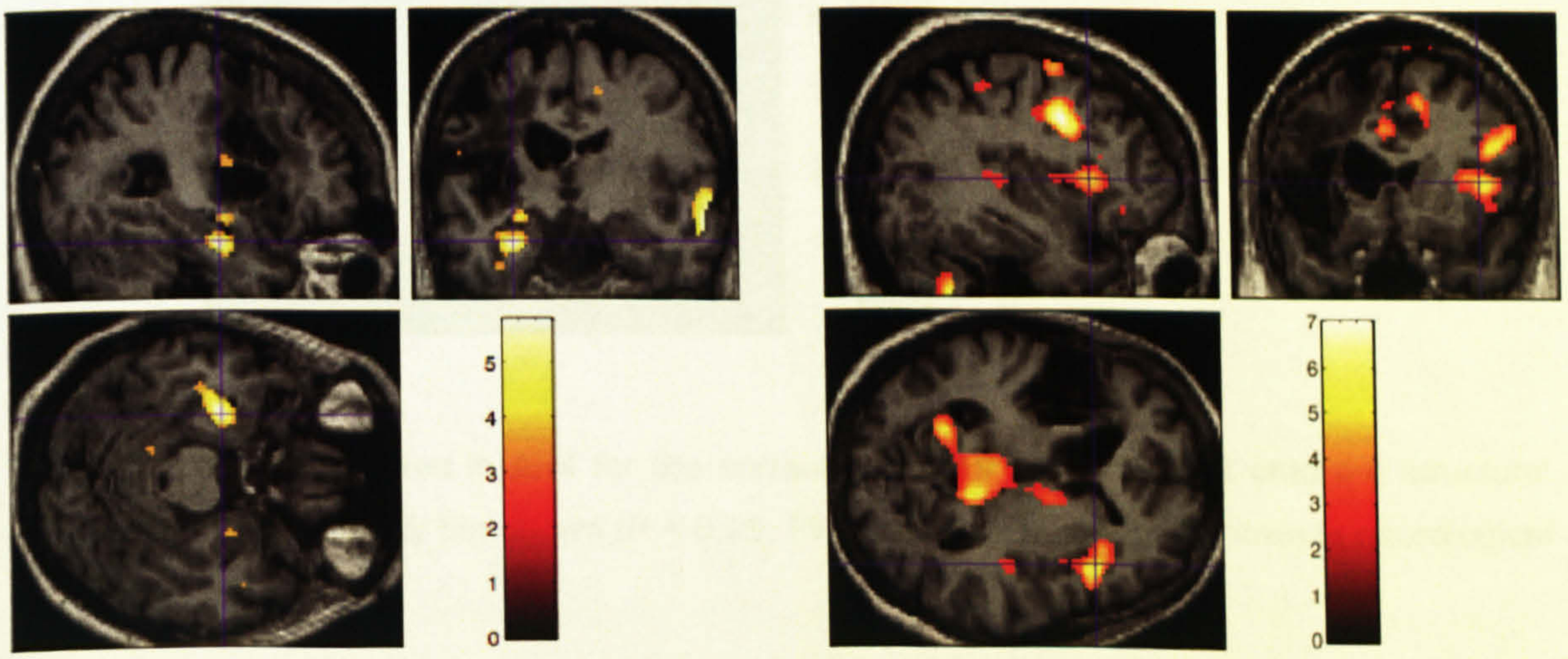

Figure 10. Areas activated in AP3 for the contrasts correct > error (left) and error > correct (right) overlaid onto the whole head volume with signal intensity bar shown (P < 0.001, uncorrected). Images shown in neurological convention, left = left.

5.5.2.4 AP4

The brain areas activated by AP4 in both correct and error naming trials are shown in Figure 11. For successful naming trials AP4 showed the largest activations in the left precentral, medial frontal and inferior frontal gyri, where Broca's area lies (BA 44). These areas were also active in the right hemisphere, though not as strongly. The right superior and middle temporal gyri showed activation, including the right hemisphere homolgue of Wernicke's area (BA 22), and left precentral gyrus and anterior cingulate gyrus. AP4's lesion covered much of Wernicke's area and occipital and temporal lobes in the left hemisphere. See Table 4 for a breakdown of all activations.

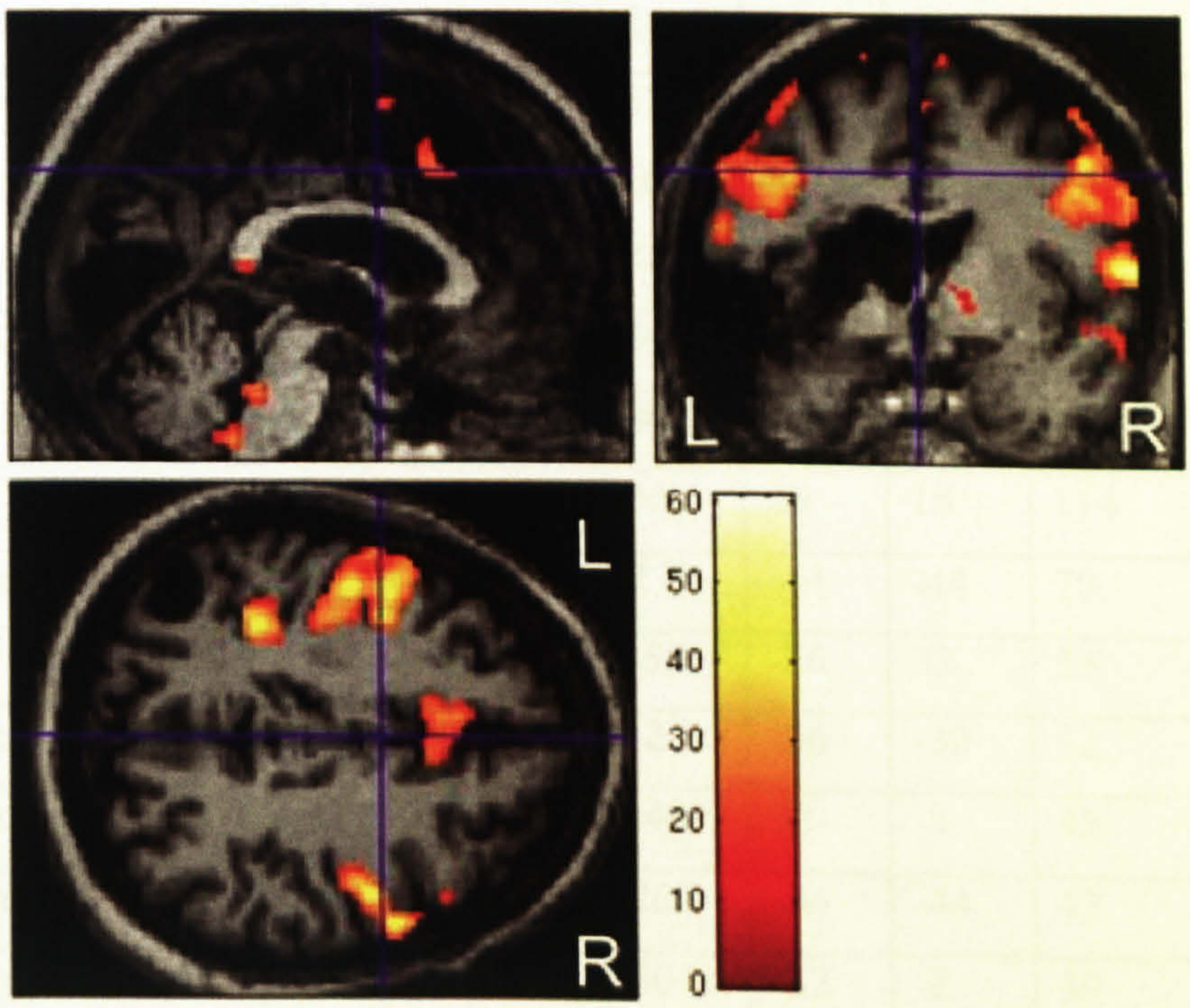

Figure 11. Areas activated in AP4 for the correct and error trials overlaid onto the structural image with signal intensity bar shown ($P < 0.05$, FWE corrected). Images shown in neurological convention, left = left.

Region	L/R	MNI coordinates			Voxels (k)	Z
		x	y	z		
Precentral Gyrus (BA 6)	L	-50	-8	42	1989	11.3
Inferior Frontal Gyrus (BA 44)	L	-54	6	22		10.7
Medial Frontal Gyrus (BA 9)	L	-48	6	28		9.6
Inferior Parietal Lobule (BA 40)	L	-34	-36	38		9.3
Precentral Gyrus (BA 6)	R	46	-6	36	688	10.7
Anterior Cingulate Gyrus (BA 32)		-6	24	36	419	6.8
Precentral Gyrus (BA 6)	L	-6	6	58		6.7
Anterior Cingulate Gyrus (BA 32)		2	16	44		6.7
Superior Temporal Gyrus (BA 22)	R	55	-36	10	370	10.6
Middle Temporal Gyrus (BA 21)	R	60	-4	-10		6.9
Heschl's Gyrus (BA 41)	R	40	-24	10		7.6
Precentral Gyrus (BA 6)	R	60	2	10	243	11.1
Inferior Frontal Gyrus (BA 47)	R	40	16	-10	133	9.6
Insula (BA 13)	R	36	20	18	114	6.3
Cerebellum	R	2	-44	-44	70	7.6
Postcentral Gyrus (BA 2)	L	-52	-26	32	53	6.5
Cerebellum	L	-2	-36	-30	52	6.7
Inferior Frontal Gyrus (BA 47)	L	-48	32	-8	45	6.1
Cerebellum	R	26	-34	-44	42	6.6
Middle Temporal Gyrus (BA 21)	R	50	-32	-2	39	6.0
Medial Frontal Gyrus (BA 9)	R	48	14	28	21	6.2
Insula (BA 13)	L	-34	-36	18	19	6.2

Table 4. Brain areas activated by AP4 during correct naming (P<0.05 FWE corrected) are shown in order of size. MNI co-ordinates are shown for activated clusters, cluster size is given in number of voxels (k) and the maximum z-score (Z) is given for each area.

Figure 12 shows the areas that saw significant differences between the correct and error naming conditions. The areas showing activation for the contrasts ((correct > error) $p > 0.001$ uncorrected) and ((error > correct), $p > 0.001$ uncorrected) are summarised in Table

5. Areas selectively active for correct naming include the right precentral gyrus, left inferior parietal lobule and right cerebellum. The cortical areas that showed more significance for the error naming trials than the correct trials include the medial frontal gyrus and insula in the left hemisphere, the precentral gyrus in the right hemisphere and the anterior cingulate gyrus.

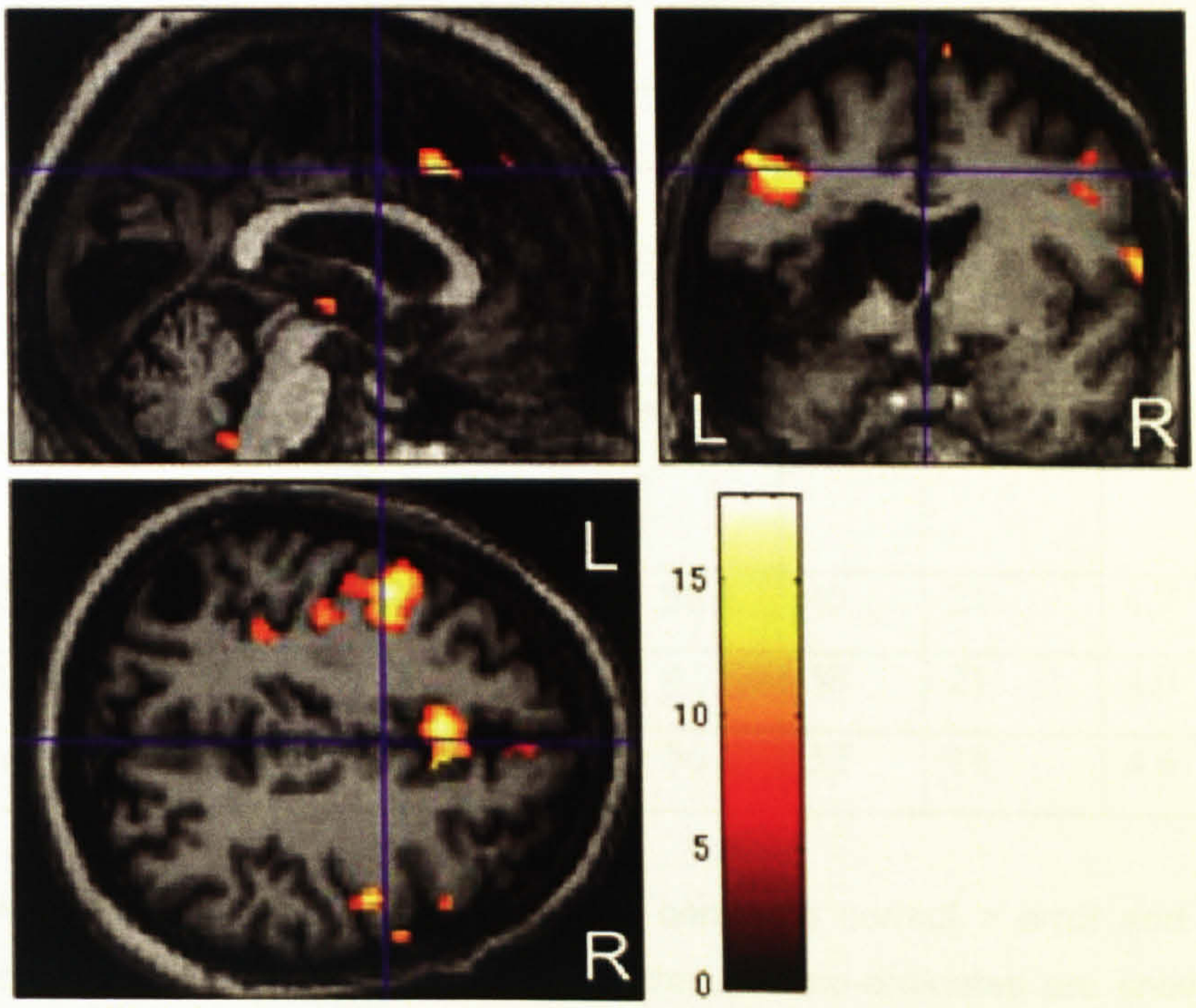

Figure 12. Areas in AP4 that show significant differences between the correct and error conditions overlaid onto the structural image with signal intensity bar shown ($P < 0.001$, uncorrected). Images shown in neurological convention, left = left.

Region	L/R	MNI coordinates			Voxels (k)	Z	P
		x	y	z			
Correct > Error							
Precentral Gyrus (BA 6)	R	48	-6	36	71	5.1	0.001
Precentral Gyrus (BA 6)	R	62	2	12	63	4.7	0.003
Inferior Parietal Lobule (BA 40)	L	-34	-36	40	33	4.0	0.009
Cerebellum	R	2	-44	-44	26	4.2	0.006
Error > Correct							
Medial Frontal Gyrus (BA 9)	L	-46	4	40	968	5.3	0.001
Insula (BA 13)	L	-40	8	20		4.8	0.002
Anterior Cingulate Gyrus (BA 32)		-4	18	38	319	5.4	<0.001
Precentral Gyrus (BA 6)	R	2	38	40	21	4.7	0.014
Precentral Gyrus (BA 6)	R	58	4	38	21	4.0	0.008
Medial Frontal Gyrus (BA 9)	L	-8	36	32	18	4.6	0.003

Table 5. Brain areas activated by AP4 for the contrasts correct > error and error > correct (P<0.001 uncorrected) are shown in order of size. MNI co-ordinates are shown for activated clusters; cluster size is given in number of voxels (k) and the maximum z-score (Z) is given for each area. The p values (P) are corrected using small volume correction.

5.5.2.5 AP5

When generating picture naming responses AP5 showed the largest activations in the anterior cingulate gyrus and the precentral gyrus bilaterally, with greater activation in the left hemisphere (see Figure 13). Activations were also seen bilaterally in the superior temporal gyrus and the inferior frontal gyrus. Activations were greater on the left for the superior temporal gyrus and on the right for the inferior frontal gyrus, homologous to the lesion site in left Broca's area. See Table 7 for a summary of all activated areas.

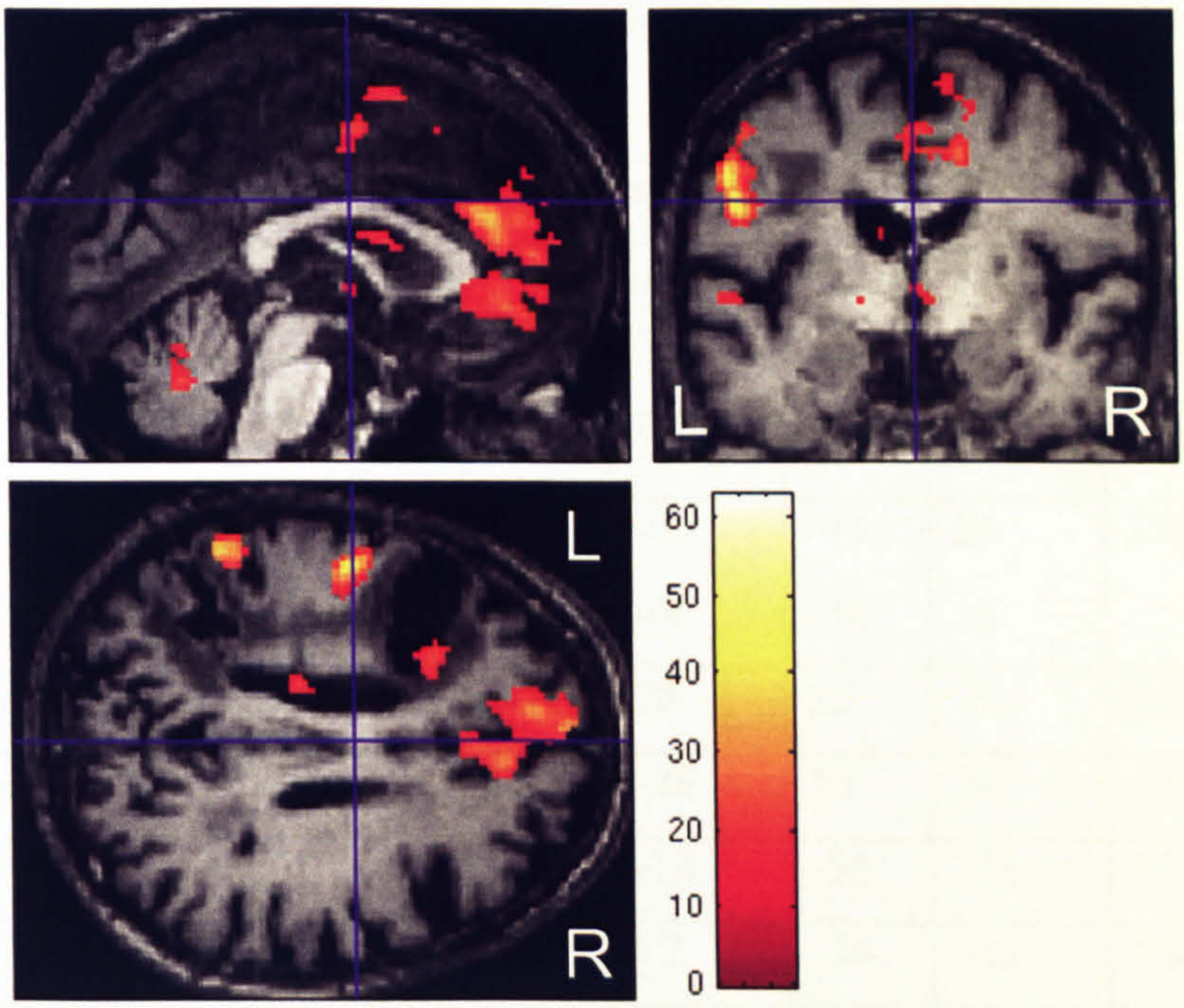

Figure 13. Areas activated in AP5 for the correct and error trials overlaid onto the structural image with signal intensity bar shown ($P < 0.05$, FWE corrected). Images shown in neurological convention, left = left.

Region	L/R	MNI coordinates			Voxels (k)	Z
		x	y	Z		
Anterior Cingulate Gyrus (BA 32)		-6	40	16	2590	11.3
Precentral Gyrus (BA 6)	L	-56	-8	36	345	10.8
Precentral Gyrus (BA 6)	R	14	-2	64	261	7.3
Inferior parietal lobule (BA 40)	L	-60	-44	26	251	9.8
Superior Temporal Gyrus (BA 22)	L	-60	-48	18		7.6
Inferior Frontal Gyrus (BA 47)	R	28	22	-8	245	6.6
Medial Frontal Gyrus (BA 9)	L	-28	16	34	181	6.5
Cerebellum	L	-12	-54	-38	172	6.7
Cerebellum	R	12	-36	-26	136	10.7
Cerebellum	R	16	-62	-20	120	7.4
Insula (BA 13)	R	32	-28	16	100	7.1
Thalamus	L	-12	-16	24	90	6.1
Precentral Gyrus (BA 4)	R	38	-20	42	82	6.2
Middle Frontal Gyrus (BA 11)	L	-36	42	-12		5.7
Thalamus	L	-16	-16	-2	69	6.7
Superior Temporal Gyrus (BA 22)	R	54	8	2	59	7.4
Middle temporal gyrus (BA 20)	R	54	-38	-14	57	7.4
Cerebellum	R	32	-40	-18	43	6.5
Superior Temporal Gyrus (BA 22)	R	62	-42	10	30	6.2
Inferior Frontal Gyrus (BA 44)	L	-50	8	6	21	6.1

Table 7. Brain areas activated by AP5 during correct naming (P<0.05 FWE corrected) are presented in order of cluster size. MNI co-ordinates are shown for activated clusters; cluster size is given in number of voxels (k) and the maximum z-score (Z) is given for each area.

Figure 14 shows areas where significant differences between the correct and error conditions were seen. No areas reached significance levels for the contrast (error > correct) and the (correct > error) contrast produced small significant increases in activation in the left inferior parietal lobule and the right precentral gyrus (see Table 8).

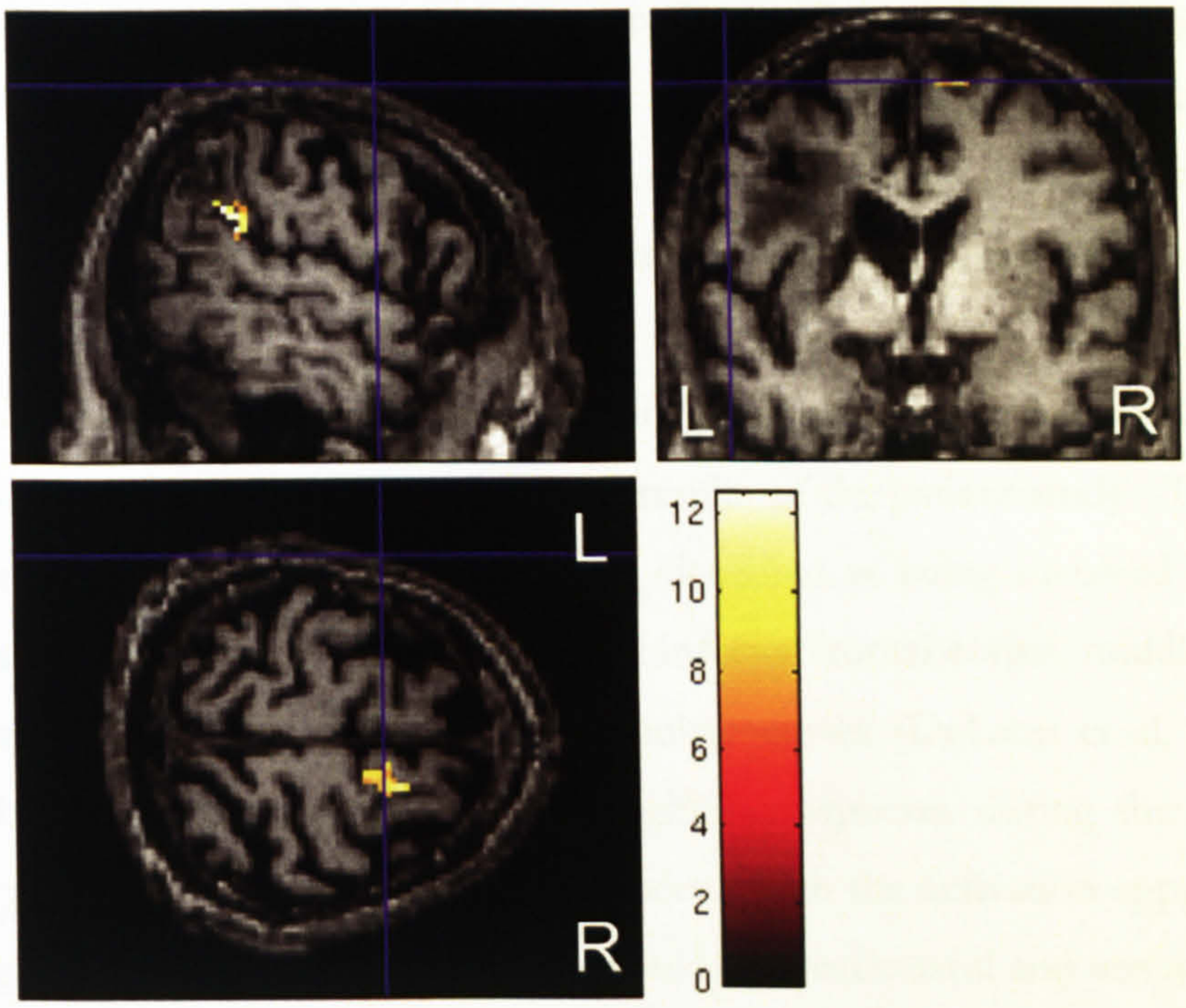

Figure 14. Areas in AP5 that show significant differences between the correct and error conditions overlaid onto the structural image with signal intensity bar shown (P < 0.001, uncorrected). Images shown in neurological convention, left = left

Region	L/R	MNI coordinates			Voxels (k)	Z	P
		X	y	Z			
Correct > Error							
Inferior Parietal Lobule (BA 40)	L	-56	-46	26	47	4.4	0.037
Precentral Gyrus (BA 6)	R	14	4	64	19	4.0	0.048

Table 8. Brain areas activated by AP5 for the contrast correct > error (P<0.001 uncorrected) are presented in order of size. MNI co-ordinates are shown for activated clusters; cluster size is given in number of voxels (k) and the maximum z-score (Z) is given for each area. The p values (P) are corrected using small volume correction.

5.6 Discussion

Five aphasic patients have undergone fMRI as part of this study. The number of potential recruits that were excluded prior to or after behavioural screening assessments had been

completed indicate the difficulty in finding patients who are able, willing and have a suitable level of aphasic deficit for inclusion in this type of study. This means that it is of real importance that functional imaging datasets acquired show valid results that are not confounded by the presence of motion during the production of task responses.

The experimental hypothesis, based on the findings from previous research, directs the focus for forming conclusions from the fMRI results of the patient study. The brain areas of interest are primarily those that have been identified as being involved in the normal processing and production of object names, i.e. inferior frontal cortex, middle and superior temporal and inferior parietal cortex and cingulate cortex (DeLeon et al, 2007). These areas may be seen active in both left and right hemispheres during the course of the experiment and the specific task conditions under which the activation appeared is crucial for inferring the cortical areas that were associated with successful and unsuccessful picture naming performance in these patients.

Particular attention needs to be paid to perilesional areas and those contralateral to the lesion site to see what differential behaviour they may display in the correct and error naming conditions. If an increase in activity is seen contralateral to the stroke damage in the left hemisphere specifically when correct, rather than error, naming responses are produced (for the contrast correct > error) then it leads to the conclusion that for this patient these right hemispheric areas can provide functionally useful language processing and may be exploited to promote recovery, supporting evidence from Basso et al (1989) and Musso et al (1999) for example. If, instead, such an increase in activation in areas homologous to the lesion occurred during error naming, but not correct naming trials (using the contrast error > correct), it would suggest that these areas were not contributing successfully to picture naming and may represent functionally inadequate or detrimental processing as found by Rosen et al (2000) and Karbe et al (1998). Such a finding would also present a focus for the development of rehabilitative therapies (Naeser et al, 2005).

The identification of cortical areas associated with successful and unsuccessful naming for each patient could inform a tailored plan of rehabilitation, enhancing the use of "correct" areas and suppressing "error" areas through the alteration of cognitive task strategies and/or the neural pattern of activation adopted to achieve naming.

The data provided by AP1 and AP2 is weak compared to that of later patients, perhaps due to the scanner equipment used at this time. Nonetheless, for both patients the areas associated with correct naming were inferior frontal and superior temporal areas in the right hemisphere, with AP1 showing a selective increase in activation in the right inferior frontal cortex and right superior temporal gyrus for correct responses as compared to error responses. AP1's lesion included the inferior frontal gyrus in the left hemisphere. The results from these two patients are contrary to the initial hypothesis and support the notion of the right hemisphere language homologues providing functionally useful processing (Blank et al, 2003).

Areas more significant for error trials compared to correct trials in AP3 included inferior frontal and premotor areas in the right hemisphere, directly opposite this patient's lesion site. Correct processing was found to be associated with increased activation in the left middle temporal gyrus, adjacent to the stroke damage. In contrast to results from the first two patients AP3's data supports previous evidence of perilesional activation being most important for functionally sound task performance, with unsuccessful processing coinciding with an increase of signal in areas of the right hemisphere homologous to the lesion (e.g. Karbe et al, 1998). These results suggest that those areas did not contribute functionally useful processing to the picture naming task and they may instead represent maladaptive cortical reorganisation following damage to the normal language network. It is interesting to note that AP3 achieved the best task performance of all the five patients. The main experimental hypothesis, that correct naming would be associated with an increase in activation in preserved left lateralised language areas and error responses would be related to processing in right hemisphere homologues, is supported by the data acquired from AP3.

AP4 showed strong activation in the anterior cingulate, left medial frontal gyrus and insula for error versus correct trials. These areas have been linked to error monitoring and performance control processes that occur in unimpaired subjects (MacDonald et al, 2000; Kan and Thompson-Schill, 2004). It had been noted that this patient had previously sometimes shown a lack of awareness of errors in speech and would neglect to correct them. It may be that at least on some level this patient was aware of errors occurring in the picture naming process, but could not correct them before the responses were produced. There was a small activation in the right precentral gyrus just for error responses, however

areas of the right precentral gyrus were also seen to be more strongly activated during correct versus error trials. During picture naming AP4 showed a largely bilateral activation pattern, including many of the areas normally expected for a naming task. Where AP4's lesion was, covering left temporal areas including Wernicke's area, strong activation was instead seen in superior and middle temporal gyri in the right hemisphere. Left Wernicke's area has been implicated in semantic processing (e.g. DeLeon et al, 2007) so it may not be surprising that of the 31 error responses AP4 produced 18 were semantic in nature. This suggests that the right hemisphere homologue to Wernicke's area may support semantic function to some degree, just not as effectively or as consistently as its damaged left counterpart. AP4's results do not go so far as to indicate that a release of right hemisphere areas from inhibition is the source of error naming, with some right precentral areas selectively involved in correct naming as well as errors. However, there is an indication that areas contralateral to the lesion site may not have the ability to fully take over the function of the damaged areas.

The results from AP5 show a large activation in the anterior cingulate that appears to be present in both response conditions. This patient's speech is very effortful and so this anterior cingulate activity may represent a high level of monitoring occurring during processing of either type of naming response, meaning that this area was active throughout. In an event-related potential (ERP) study of action monitoring Gehring & Knight (2000) found that individuals with damage to the lateral prefrontal cortex, as AP5 shows, produced equal anterior cingulate response to error and correct trials in contrast to controls, who showed the normal error related increase. They concluded that the lateral prefrontal cortex normally interacts with the anterior cingulate in action monitoring and may provide information to aid discrimination of correct and error responses. AP5's lesion in this area may be giving rise to the consistent anterior cingulate activation seen in this study. AP5's results do suggest that anterior cingulate activity alone is not enough to implement the correction of errors. The specific role of the anterior cingulate in the presence of errors is not fully understood however and is the subject of ongoing research.

Returning to the issue of the role of the right hemisphere in aphasic naming performance, no areas in either the left or right hemispheres showed selective activation for error responses in AP5. A small but significant activation in the right precentral gyrus was recorded in the correct versus error condition, meaning that this area was more strongly

activated during correct naming performance than incorrect performance. The right inferior frontal gyrus showed much greater activation than the damaged left inferior frontal gyrus during the task, however, there was no significant difference in activation levels between the two response conditions. This may indicate that the right inferior frontal gyrus was providing some functional support for its lesioned left homologue, but just was not as efficient. AP5's data, however, does not provide evidence for the specific contribution of right hemisphere areas to error production.

Considering all single subject data from the aphasic patients together it presents a mixed picture, emphasising the influence of individual differences and the difficulty of studying groups extracted from the aphasic population who often have widely varying language and related stroke induced deficits. On the whole, the results do not support the initial hypothesis that correct naming performance would be associated with processing within preserved language areas in the left hemisphere whereas the production of errors would coincide with an increase in activation in right hemisphere language homologues. Instead, results from AP1 and AP2 support the reverse of this hypothesis by showing a selective increase in activation in right hemisphere areas homologous to their lesion site for correct trials, suggesting that these areas can promote functional recovery from aphasia.

AP4 and AP5 still show a largely bilateral activation pattern in response to the naming task, with the left hemisphere more dominant. For these two patients the right hemisphere areas homologous to their lesions were not seen to be active when the contrasts correct > error or error > correct were applied, but were active when correct and error trials were considered together. This suggests that there was no significant difference in extent or strength of activation in these areas between the two conditions. The data cannot tell us whether these lesion homologues in the right hemisphere were activating more strongly than they would normally in the pre-damage language system and so are part of reorganisation of processing, however, given that these patients still both made a substantial amount of errors it can be concluded that these areas were not sufficiently taking over function from their damaged counterparts.

It is only AP3 who provides some support for the hypothesis, showing an increase in activation in the right inferior frontal cortex and right premotor area for the error > correct contrast and perilesional activation for the correct > error calculation. AP3's lesion

included Broca's area in the inferior frontal cortex and premotor area in the left hemisphere. This patient's results clearly show that activating the normal left lateralised language system around the damaged cortex is associated with successful picture naming performance, whereas errors coincided with an increase in signal in areas homologous to the lesion. These results agree with previous work that have led to suggestions that reorganisation of language processing to right hemisphere homologues after damage to previously dominant areas of the language system may be more interfering than compensatory (e.g. Rosen et al, 2000; Naeser et al, 2005). However, it has to be reiterated that this result is only from one patient in a series of five single subject analyses and this particular patient suffered the sharpest movements when speaking their response.

Most of all these results support the conclusion that under certain circumstances either perilesional or contralateral reorganisation of function can contribute to recovery from aphasia (Crosson et al, 2007). The data from this study points towards shifting the research focus to identifying the patients who would benefit from reorganisation of cortical or cognitive processing to activate alternative pathways and those who should be helped to reactivate perilesional sites. The small number of individual patients studied within this project prevents the recognition of clear patterns of patient characteristics that could be associated with either recovery strategy. However, it can be said that the patient whose correct naming responses were accompanied by left temporal activation, adjacent to the stroke damage, produced the best task performance. This echoes earlier findings of an association between good language recovery and reactivation of the normal left lateralised language brain areas, with worse recovery from aphasia seen in patients who persist in showing increased levels of right hemisphere activation (e.g. Karbe et al, 1998; Rosen et al, 2000; Perani et al, 2003). Extending the study by acquiring images from more people with aphasia would provide a greater opportunity to search for connections between specific patient and lesion characteristics and the degree of recovery of language function attained.

Turning attention to the methodological issues highlighted by this study, the problem with motion artifact seen with AP3 underlines the need for further research into effectively reducing the impact of motion related to overt speech without compromising the data. The variability of speech onset times displayed between patients and between trial types was accounted for in this study by including exact response times in the general linear model. This is allowed for when images are acquired continuously as in this study,

however, it is difficult to see how this issue could be effectively dealt with using a sparse paradigm that only samples certain time points within a trial. Sparse paradigms have recently become a common way of allowing for spoken responses within the fMRI scanner whilst sidestepping the associated motion problems (Fridriksson et al, 2006; Abrahams et al, 2003). These results question whether these studies are losing valuable data by adopting this method.

In conclusion, this study was successful at applying an event-related, continuous scanning protocol to the study of correct and error naming responses in post-stroke patients who have aphasia, fulfilling one of the initial aims of the research project. The picture naming paradigm was successfully completed by all five patients who progressed to the scanning stage of the project, despite their individual variations in aphasic characteristics and lesion location. A reasonable ratio of correct to error responses was recorded from each patient, allowing for valid statistical comparison of the two trial types and indicating that the adopted screening process was appropriate for identifying the degree of anomic deficit sufficient for this study. The theory of right hemisphere language homologues contributing to the production of naming errors when the normally dominant left hemisphere is lesioned was not fully supported by these results, contrary to the initial hypothesis. Only data from one patient could provide evidence for this conclusion. Results from the four additional patients suggest a positive role for contralesional brain areas in the post-stroke recovery of naming ability, albeit with limited effectiveness. Implications of these results for therapeutic practice and the planning of future studies of recovery from aphasia are considered in the general discussion (Chapter 7).

6 Forcing Errors in Unimpaired Systems

6.1 Did I Really Say That?

Speech errors are not exclusive to people with an impaired language system. Everyone will be familiar with the frustration of the "tip of the tongue" state, where the desired word remains stubbornly elusive to the speaker's lexical selection process. Embarrassment at spontaneous substitution of a congruous word in a sentence with an inappropriate replacement or neologism can seem all too commonplace, especially when tired or distracted. It has been suggested that normal and aphasic speech errors exist on a continuum, with an aphasic language performance representing an exaggeration of normal impairments that occur when the language network is not running optimally (Silkes et al, 2004). Speech errors occurring normally are said to be qualitatively of a similar nature to those produced by aphasics whose language system has been physically damaged, with the major differences lying in quantity and severity of the errors. This has been referred to as the "continuity thesis" by Dell et al (1997). If this is true then can studies of speech errors in unimpaired speakers inform knowledge of the mechanisms of such errors in aphasics?

6.1.1 Models of normal and abnormal language processing

Accounts of the deficits suffered by post-stroke aphasic patients can be generally separated into two groups according to the mechanisms that are theorised to be responsible for the reduction in linguistic ability. Some theorists believe that a lesion to a brain region involved in the functional language network will result in loss of information, or a specific linguistic ability, that was stored in, or connected to, that particular cortical area. In this case an aphasic deficit is thought to directly map onto the tissue damage incurred through brain injury. This kind of language model tends to be formed from a processing stream of sequentially organised language components. This type of account, however, does not attempt to explain how errors in an undamaged system occur. It regards aphasia as a set of specific conditions in which the handling of language comprehension and/or production has been forced to deviate significantly from what is normal.

The common alternative view focuses more on aphasia as a deficit in system-wide processing. Instead of the lesion resulting in the loss of linguistic information, impaired performance is linked to a reduction in ability to access or activate the appropriate language elements (Silkes et al, 2004). Following on from this interpretation of aphasia psycholinguistic models have been developed that attempt to explain aphasic and normal speech within the same language processing architecture. These models are typically designed as interactive parallel processing units where spreading activation is used to stimulate target nodes in a network (Bock and Levelt, 1994; Dell et al, 2007; Schwartz et al, 2006). Figure 1 presents an example diagram of an account of object naming using spreading activation.

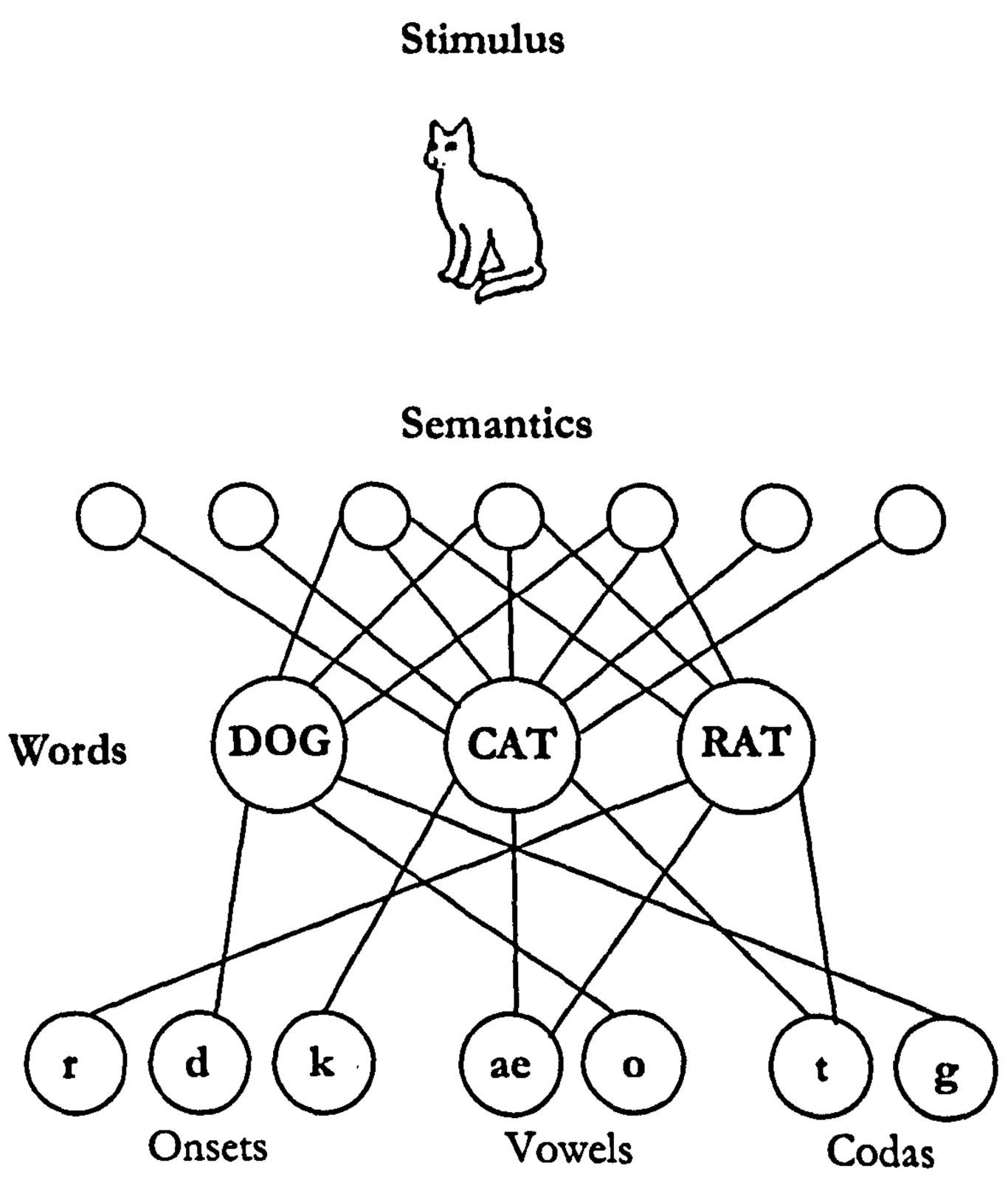

Figure 1. Example of a spreading activation account of the object naming process.

The connectionist models typically begin with an account of unimpaired processing and then "lesion" the model by altering certain parameter values in some quantifiable way to

simulate the effects of brain damage on the language system. In this way the deliverance of activation to the target node(s) is affected, making errors more likely. Examples of this can be found in the study of comprehension (Haarmann et al, 1997), dyslexia (Plaut et al, 1996) and naming (Foygel and Dell, 2000). Recovery is then characterised as the return of disordered parameter values towards the normal functioning state.

This view that aphasic errors result from adverse alterations to parameter values within the normal language system is shared by Schwartz et al (1994). Whereas Dell et al (1997, 2007) seek to provide evidence in the form of developing a computer model that fits the data from a corpus of normal and aphasic naming errors, Schwartz et al take a different approach. They conducted a study using tongue twisters to attempt to show that aphasic and normal errors exist on the same continuum and can therefore arise from the same mechanisms. Nonaphasic subjects were given a set of difficult tongue twisters to say for a total of sixteen times each during the experiment. The experimenters were primarily interested in the effect of practice on the number and nature of the errors produced. Within the context of the spreading activation model practice is theorised to strengthen the connections between nodes and so improve task performance and reduce errors.

The results showed that the errors produced at the outset of the task were plentiful and reproduced the main characteristics of a jargon aphasic's error pattern that they had previously analysed. As expected, the number of errors was reduced through practice, but it was also found that the pattern of errors moved from being close to the output of the jargon aphasic towards the error pattern gleaned from a corpus of normal errors. Schwartz et al characterised this as a shift from a "bad" error pattern to a "good" one, consisting of less perseveratory errors, when responses to previous stimuli are given erroneously, more anticipations and an increase in errors forming recognisable words rather than neologisms. This change in the nature of errors made is claimed to reflect the level of efficiency in the language system, where efficiency is defined as how successfully activation is delivered between units within the time that is available to complete the operation. An efficient system given the task of naming a picture should activate target nodes strongly enough to overcome any residual activation from previous stimuli, reducing the likelihood for perseverations. Connecting nodes will also receive some activation, meaning that any errors made should be biased towards creating other known words. An efficient language system can be compromised by a weakening of the connections due to physical damage or

more normal factors such as lack of practice with novel words or strings, tiredness or increasing the rate of speech.

6.1.2 Inducing errors in the normal system

Limiting the amount of time available to produce a response is an effective way of creating or enhancing weaknesses in a language system that relies on spreading activation. It can prevent activation from being fully delivered to the target destination, thus reducing the dominance over competing nodes and increasing the interfering effect from previously activated nodes. The likelihood of a spreading activation system delivering a "bad" error pattern when given a language task to complete is increased. This theoretical outcome of disrupting the behaviour of the normal language system in this way has been used by a number of researchers to attempt to simulate paraphasias in unimpaired speakers under experimental conditions.

Vitkovitch and Humphreys (1991) imposed a response deadline of 600 msec on participants naming pictures with high and low frequency names. This forced their subjects into making more errors than in naming studies without a response deadline. Pictures with low frequency names induced more errors than those with high frequency names. The results showed that up to 20% of responses to the low frequency names were errors; this was compared to an error rate of less than 10% in equivalent studies without a deadline. Vitkovitch and Humphreys argued that the influence of name frequency on the error rate in the task signified that the response deadline constrains the name retrieval stage of the naming process. If the introduction of the deadline affected any other stage in the process, such as visual encoding or semantic access, then errors would reflect this and should not show a name frequency effect.

A 600 msec naming deadline was also used by Laws (2000) in their study of category specific naming errors. The naming-to-deadline paradigm was utilised with normal participants in order to constrain the name retrieval process and increase the number of errors produced. Laws set out to test the predictions made by theories of category-specific deficits in aphasia by assessing whether the naming problem represents an exaggeration of normal tendencies towards one semantic category over another. Most models of category-specific deficits in aphasia imply that such effects will also appear in normal processing,

with aphasic performance largely representing a quantitative rather than qualitative difference. Laws took advantage of the error rate induced by this task to make comparisons between the responses of their normal participants and the performance of aphasics reported in the literature. The type of errors produced by their normal subjects led Laws to conclude, in agreement with Vitkovitch and Humphreys (1991), that naming difficulties induced by the naming-to-deadline task arise at the lexical processing level. This conclusion was based on the correlation between errors with the number of alternative names possible for the stimulus item and the age of acquisition of the name. It was posited that in a linguistic account the influence of these factors on naming may reflect the reduction of activation arriving at the lexicon from the semantic system.

In their study of perseverative errors, where an earlier response to a sequential task is erroneously repeated, Moses et al (2004) employed the naming-to-deadline task to force unimpaired subjects to increase their error rate. They noted that there was evidence present in the literature of perseverative errors existing in healthy speech, just at a reduced rate than in aphasic speech. Artificially increasing the number of perseverative errors made by normal subjects may prove to be a valuable method of studying the underlying nature of this error type. The reasoning for adopting this particular paradigm design in their study included the relative methodological ease of studying healthy participants rather than a heterogeneous group of aphasic speakers. Specific language impairments rarely exist alone in the aphasic condition and so removing the chance of such confounding factors proved to be another desirable design decision. The results from this study are also interpreted within the boundaries of a spreading activation linguistic account of normal and abnormal picture naming.

Silkes et al (2004) chose to manipulate the temporal variable in their picture naming task performed by non-brain damaged young adults. No actual deadline was imposed but participants were instructed that their reaction times were critically important and they should concentrate on the speed of their response rather than their accuracy. At various points during the experiment the experimenter would urge the subjects to respond more quickly to stimuli. Silkes at al reasoned logically that if an aphasic's linguistic processing is inefficient it would also take more time to complete the necessary processing for successful task performance. In this way decreasing the available time for non-brain damaged subjects was thought to be one way to simulate an aphasic naming deficit. Their results

showed that the time stressed participants generated more errors (28.7% error) than controls who completed the picture naming task in their own time (6.7% error), but less than aphasic participants they were compared with (48.8%).

Evidence presented here clearly shows that limiting the amount of time available to healthy, unimpaired participants to complete a language task increases the likelihood of errors. Therefore, the naming-to-deadline paradigm is a valid method for inducing a larger number of errors than would otherwise occur, for the purpose of statistically comparing correct and error naming trials in a functional imaging study of the normal language system. It should be acknowledged here that simply making the assumption of artificially induced errors being equal in terms of cognitive processing to aphasic errors may be a mistake. However, even if this is not so, studying naming errors made by a normal language system using functional imaging can give insights into brain areas associated with the production and monitoring of normal naming errors. This could provide a more suitable baseline for comparison with aphasic naming errors than largely correct normal naming performances.

6.2 Aims and Hypotheses

This study was conducted in order to compare brain areas activated during correct and induced error naming in unimpaired speakers with the correct and error naming trials of aphasic patients. If the normal language system were overloaded by limiting the time available to respond, would it reduce inhibitory power over non-dominant areas as seen after brain damage? If this connectionist theory is correct then studying errors forced from an undamaged language system may produce useful conclusions that could be applied to the aphasic population. Since it is practically and technically much more difficult to conduct viable fMRI experiments with stroke patients, as discussed in Chapter 2, an ability to substitute this subject group with healthy participants, who are more robust, plentiful and experimentally less complicated, would be very advantageous. This method could prove to be useful when investigating language recovery within the brain, therefore, this study aims to test this assumption.

A further interest in conducting this study lies in identifying the cortical areas involved in the processing of confrontation naming errors by an undamaged language system. In the

standard confrontation naming paradigm unimpaired participants make very few errors, as demonstrated in the two pilot studies conducted as part of this research project. Such ceiling level performance does not allow for any exploration of the neural behaviour surrounding the production of naming errors when no damage is present. It would be theoretically interesting to compare the patterns of activation during normal naming error production to those of aphasic naming errors.

It is predicted that placing the healthy participants under the pressure of a temporal deadline will significantly increase the number of errors made in a picture-naming task. Connectionist models of normal and abnormal language lead to the hypothesis that the forced error task will move participants closer to an aphasic pattern of performance. It is also expected that cortical activation patterns will include areas associated with the normal control and monitoring of performance, i.e. anterior cingulate cortex and inferior frontal cortex (Fu et al, 2002; Christoffels et al, 2007).

6.3 Pilot Study

To allow for a better match to the activation paradigm and scanning protocol already established for overt naming tasks within this project, the task, stimuli and experimental timings were kept as similar to the patient study as possible. Although previous naming-to-deadline studies show that this is an effective method for increasing normal error rates no study used the exact parameters to be used here. Therefore, prior to the fMRI study of forced errors a behavioural pilot study was carried out to ensure that the paradigm to be used in the scanning study was indeed likely to result in an increase in the production of naming errors in healthy participants.

6.3.1 Subjects

Thirteen unimpaired speakers with no neurological history were recruited for this pilot study. Informed consent was obtained and ethical approval was provided by the University of Nottingham Psychology Department ethics committee. The subject group consisted of university undergraduates with an age range of 18-22, mean age 19.15 years.

6.3.2 Stimuli

In line with all other experiments conducted as part of this project the stimuli used in this pilot study were from the standardised stimulus set of Snodgrass and Vanderwart (1980). The pictures consisted of black and white line drawings selected from across the range of the ten major semantic categories included in the stimulus set. Eighty pictures were presented to the participants for speeded naming.

6.3.3 Experimental procedure

The experimental procedure was modelled on that used by Vitkovitch and Humphreys (1991) in their speeded picture naming study. They chose a deadline of 600 msec based on the reaction times recorded in their previous naming experiments (Humphreys et al, 1988). In these experiments few participants responded quicker than 600 msec and so it was thought that this deadline would exert enough pressure on subjects to induce errors, as the results of their speeded naming study did indeed demonstrate.

The selection of this time is also supported by work done by Levelt et al (1998) on the timecourse of the naming process. Collating results from previous MEG and ERP studies they defined the time window for the last stage of naming, phonetic and articulatory processing, as 400-600 msec. Therefore, setting the verbalised response deadline at 600 msec should restrict the time subjects would normally spend naming a picture.

In this forced error pilot study each stimulus picture was displayed on a laptop computer for 600 msec. At 600 msec a buzzer sounded, also for 600 msec, and the picture was removed (see Figure 2). The removal of the picture served as an additional deadline cue and prevented the participants from continuing to process the stimulus item in their own time before producing a response. Participants were instructed to name each picture as quickly as they could and to try and "beat the buzzer". They were advised to concentrate on the speed of their response rather than the accuracy. Each trial was 15s long to replicate the trial time used in the fMRI study of aphasic patients. There were 80 trials, giving a total experimental time of 20 minutes. Responses were recorded by the laptop computer using a microphone and were transcribed by the experimenter during the experiment. Transcriptions were later checked with the recorded output and reaction times extracted from the sound file.

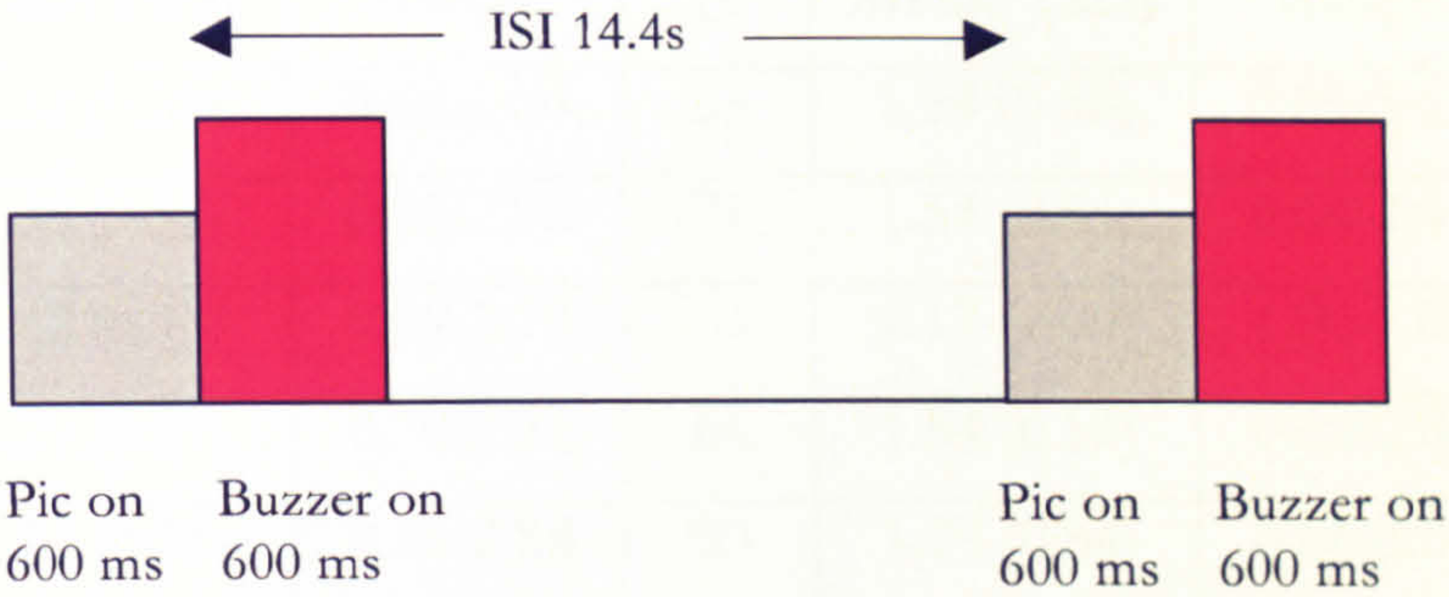

Figure 2. Paradigm used for the forced error behavioural pilot and subsequent fMRI study.

Responses were categorised using the same classification system as for the patient study (see Chapter 5, section 5.4.3.1) where responses were coded as correct or into one of the error categories of semantic, formal, mixed, unrelated, non-word and no response (Dell et al, 1997; Caramazza et al, 2000; Capitani & Laiacona, 2004). As before, the categories semantic/visual and unrelated/visual were also used when the target and the given error response were visually very similar. Some responses could not be precisely identified and classified, although it was clear that these responses did not match the target name or its synonyms. These errors were included in the additional category indistinguishable.

6.3.4 Results

Participants achieved a mean naming accuracy of 71.44%, which equates to 57.15 correct and 22.85 error trials from a total of 80. Table 1 shows the mean, standard deviation and range of response times of all 13 participants and the number of correct and error trials. Four participants managed to beat the buzzer on at least one occasion to produce a naming response before 600 msec, but only when that response was correct. No participants were quicker than 600 msec when generating an error and no participant achieved a mean response of less than 1 second for either condition.

	Correct Responses			Error Responses		
	Mean (SD)	Range	N	Mean (SD)	Range	N
P1	1.30 (0.38)	0.58-2.00	53	1.54 (0.58)	0.85-3.05	27
P2	1.38 (0.56)	0.55-2.98	71	1.54 (0.64)	0.64-2.43	9
P3	1.42 (0.45)	0.69-2.79	64	2.17 (1.00)	1.00-5.01	16
P4	1.29 (0.37)	0.70-2.22	60	1.54 (0.54)	0.62-2.83	20
P5	1.57 (0.56)	0.66-2.88	53	1.79 (0.84)	0.66-3.55	27
P6	1.44 (0.39)	0.65-2.46	59	1.74 (0.66)	0.60-3.01	21
P7	1.40 (0.48)	0.62-2.77	51	1.46 (0.55)	0.71-3.13	29
P8	1.67 (0.54)	0.57-2.75	44	1.72 (0.70)	0.63-3.58	36
P9	2.28 (0.83)	1.00-6.72	59	3.04 (1.28)	1.29-6.83	21
P10	1.43 (0.52)	0.67-3.36	51	1.72 (0.71)	0.67-3.27	29
P11	2.04 (1.43)	0.57-8.72	63	1.79 (0.87)	0.82-3.89	17
P12	1.45 (0.51)	0.67-3.05	53	1.43 (0.42)	1.00-2.55	27
P13	1.30 (0.37)	0.61-2.62	62	1.76 (0.68)	0.90-3.28	18

Table 1. Speech onset times for correct and error naming responses measured in seconds and the number of corrects and errors produced by each participant out of eighty trials.

Table 2 shows the distribution of errors across the different error categories for all the participants together. Errors produced were predominantly semantic in nature, accounting for 44% of the total number of errors with a further 11% being both semantically and visually related to the target name. The semantic categories were the only error types to be produced by all participants. The category of no responses, or omissions, was the next most represented error, with a 19% share. This result in the healthy subjects is in line with previous studies of speeded naming that also found semantic errors to be the dominant type induced by a response deadline (Vitkovitch & Humphreys, 1991; Moses et al, 2004; Laws, 2000).

Error Type	Mean (SD)	Range	%
Semantic	10.08 (4.23)	3 – 20	44.11
Semantic/Visual	2.46 (0.88)	1 – 4	10.77
Formal	0.38 (0.77)	0 – 2	1.68
Mixed	1.62 (1.19)	0 – 4	7.07
Unrelated	0.85 (1.07)	0 – 3	3.70
Unrelated/Visual	0.38 (0.65)	0 – 2	1.68
Non-word	0.23 (0.44)	0 – 1	1.01
No Response	4.38 (5.39)	0 – 15	19.19
Indistinguishable	0.23 (0.44)	0 – 1	1.01
Other	2.23 (2.39)	0 – 8	9.76

Table 2. Mean and range of number of errors in each error category across all participants and the percentage share of error types.

It was noted, anecdotally, that a small number of participants would cease to attempt to beat the buzzer with their response as the experiment progressed, resulting in longer reaction times and a reduction of their error rate relative to other participants. Also, some subjects tended to refrain from producing a response if they were unsure of the correct name, leading to a high number of no response trials. These observations underlined the need to emphasise instructions for the fMRI study to try and produce a speeded response and to give a response wherever possible, ignoring its accuracy.

These results show that imposing a deadline of 600 msec does exert enough pressure on the normal naming process to induce an increased error rate. The forced error pilot participants achieved a mean of 29.66% errors compared to that of only 3% errors made by the participants in the original fMRI naming pilot study (see Chapter 4), although the two studies were conducted in different environments and with a difference in the amount of naming stimuli which may have contributed to the variance in error rate. The results support those of Vitkovitch and Humphreys (1991) who, using the same naming deadline, found that naming errors increased up to 20% compared to rates of less than 10% in unconstrained naming studies. The error rate of forced error participants is also very similar to that found by Silkes et al (2004), where their non-brain damaged participants completing a naming task under time pressure produced an error rate of 28.7%. This was a

significant increase from the 6.7% of errors made by controls who were not working under any time constraints.

6.4 Forced Error fMRI Study

6.4.1 Subjects

The subjects recruited for the functional imaging study had not previously participated in the pilot study. 8 university undergraduates were tested, aged between 18 and 20. Handedness was assessed with the Edinburgh Handedness Inventory (Oldfield, 1971) and all subjects were found to be right-handed. Written consent was sought from all taking part and it was ensured that there was no reason for subjects to be prevented from being scanned, such as the presence of metallic implants.

6.4.2 Activation paradigm

The paradigm used in the fMRI study was identical to that of the behavioural pilot study. The same set of eighty Snodgrass and Vanderwart (1980) pictures were presented in a randomised order for naming. Participants viewed the pictures via Avotech "Silent Vision" eye pieces (http://www.avotec.org) that were mounted on to the scanner's head coil and responses were recorded using the same microphone setup as used in the fMRI pilot study (see Chapter 4), based on the system described by Barch et al (1999). As before, stimuli were shown for 600 msec with a buzzer sounding at this time to coincide with the removal of the picture. A fixation cross was shown in the centre of the display during the 14.4 sec inter-stimulus interval (ISI) to act as a guide to the placement of subsequent pictures. This paradigm was controlled by the software Presentation (http://neurobs.com). Scan time for the eighty naming trials was twenty minutes.

6.4.3 Scanning parameters

T2* weighted EPI images were acquired using a Philips 3T Achieva fMRI scanner with an 8-channel SENSE head coil. 18 coronal slices were imaged to cover the field of view of the posterior frontal and temporal lobes (see figure 8 in Chapter 4 for a diagram showing the field of view) with a resolution of 3 x 3mm in-plane, 6mm slice thickness, TE = 35ms, sense factor = 2. 454 volumes were captured for each participant with a repetition time

(TR) of 2.00s. These parameters are identical to those used in the study with aphasic patients.

6.4.4 Data processing and analysis

6.4.4.1 Spoken responses

All digital recordings were passed through a noise cancellation tool (Cusack et al, 2005) to reduce the level of scanner noise interference from the sound file. Responses were then transcribed and categorised as either correct or error for the purpose of separating fMRI trials into the two conditions post-hoc. Exact verbal response onset times were recorded for use when building the general linear model design matrix in SPM5. Error responses were further split into the different error types using the classification system employed previously in this project (see Chapter 5, section 5.4.3.1) in line with previous naming studies (Dell et al, 1997; Caramazza et al, 2000; Capitani & Laiacona, 2004). This was done for behavioural analysis only as it was not expected that there would be enough trials of each error type for inclusion in the statistical analysis of the fMRI data.

6.4.4.2 fMRI data

As per the patient study presented in the previous chapter the fMRI data was pre-processed and statistically analysed using SPM5 (http://www.fil.ion.ucl.ac.uk/spm/). Chapter 4 contains descriptions of the pre-processing and analysis methods used in subsequent chapters.

Images for the young forced error group were slice timed and motion corrected before being normalised to MNI space. Spatial smoothing was applied to the data using a gaussian kernel of 8mm and temporally filtered with a 30s cut-off point specified. Events were defined in the statistical model as either correct or error, using the exact response times to specify the onset of the stimulation with a duration of 1s.

Data from these subjects was combined as a group and entered into a fixed effects analysis. This analysis was performed in the same way as the individual patient statistics detailed in Chapter 5. Activation maps were produced for the correct and error naming response conditions ($p < 0.05$, using family-wise error (FWE) correction for multiple comparisons). A statistical map of all activation in both correct and error conditions was generated using a

threshold of $p < 0.001$ uncorrected. The resulting image was saved for use as a masking image in a small volume correction to be applied to the contrasts of conditions. This correction reduced the search volume to decrease the number of statistical tests calculated across the dataset and to limit the number of false positives likely to be present in the resulting activation map.

Significant activations detected during correct naming and the production of errors were directly compared using the contrasts error > correct and correct > error ($p < 0.001$, uncorrected, masked by an inclusive mask of (error and correct) conditions at $p < 0.001$, uncorrected) to show areas selectively activated for each type of naming response.

6.4.5 Results

6.4.5.1 Naming performance

The young forced error group produced a mean number of 53.63 correct responses (SD = 5.58) and 26 errors (SD = 5.07) out of a total of 80 trials (see Table 3). Participants tended to produce quicker responses when retrieving a correct name. Only one participant managed to beat the deadline on at least one occasion in error trials, whereas five participants achieved response times faster than 600 msec for the correct condition.

	Correct Responses			Error Responses		
	Mean (SD)	Range	N	Mean (SD)	Range	N
YP1	1.59 (1.29)	0.74-9.38	47	2.20 (1.07)	0.99-4.94	33
YP2	0.68 (0.09)	0.53-0.89	46	0.72 (0.08)	0.60-0.83	31
YP3	0.77 (0.16)	0.54-1.56	54	0.79 (0.14)	0.63-1.14	26
YP4	0.79 (0.23)	0.54-1.60	62	1.34 (0.95)	0.64-3.34	18
YP5	0.90 (0.23)	0.61-1.92	60	0.92 (0.22)	0.69-1.46	20
YP6	0.95 (0.50)	0.53-3.05	55	0.91 (0.48)	0.61-2.62	25
YP7	0.87 (0.18)	0.63-1.57	52	0.99 (0.38)	0.54-2.02	28
YP8	0.78 (0.11)	0.59-1.08	53	0.87 (0.09)	0.73-1.08	27

Table 3. Speech onset times for correct and error naming responses measured in seconds and the number of corrects and errors produced by each young participant.

Table 4 shows that the predominant error types for the young group were semantic errors, with a 34.13% share, and no responses (35.10% share). No participants made a non-word error and very few errors of a formal nature were produced. Some responses were indistinguishable due to the spoken response coinciding with the buzzer sound, although it could be distinguished that these responses did not match the target picture name.

Error Type	YP1	YP2	YP3	YP4	YP5	YP6	YP7	YP8	%
Semantic	14	9	7	5	8	13	7	8	34.13
Semantic/Visual	2	0	1	1	2	1	0	3	4.81
Formal	0	0	1	0	0	1	0	0	0.96
Mixed	1	1	0	0	0	1	0	2	2.40
Unrelated	1	2	2	2	1	0	4	0	5.77
Unrelated/Visual	0	0	0	1	0	1	0	0	0.96
Non-word	0	0	0	0	0	0	0	0	0
No response	4	16	10	7	7	1	16	12	35.10
Indistinguishable	0	3	5	1	1	3	0	2	7.21
Other	11	0	0	1	1	4	1	0	8.65

Table 4. Breakdown of the number of different error types made by the young participants, with the percentage of error trials represented by each error type.

6.4.5.2 fMRI results

Figure 3 shows the cortical areas that show differences between correct and error responses in the young forced error group. The areas are summarised in Table 5. Correct responses are associated with large activations in the anterior cingulate gyrus, accompanied by activations in the superior temporal gyrus bilaterally, the supermarginal gyrus, superior temporal gyrus and inferior frontal cortex all in the left hemisphere. In comparison, few areas were selectively active for error trials. Those areas were the anterior cingulate gyrus and left middle and inferior frontal gyri.

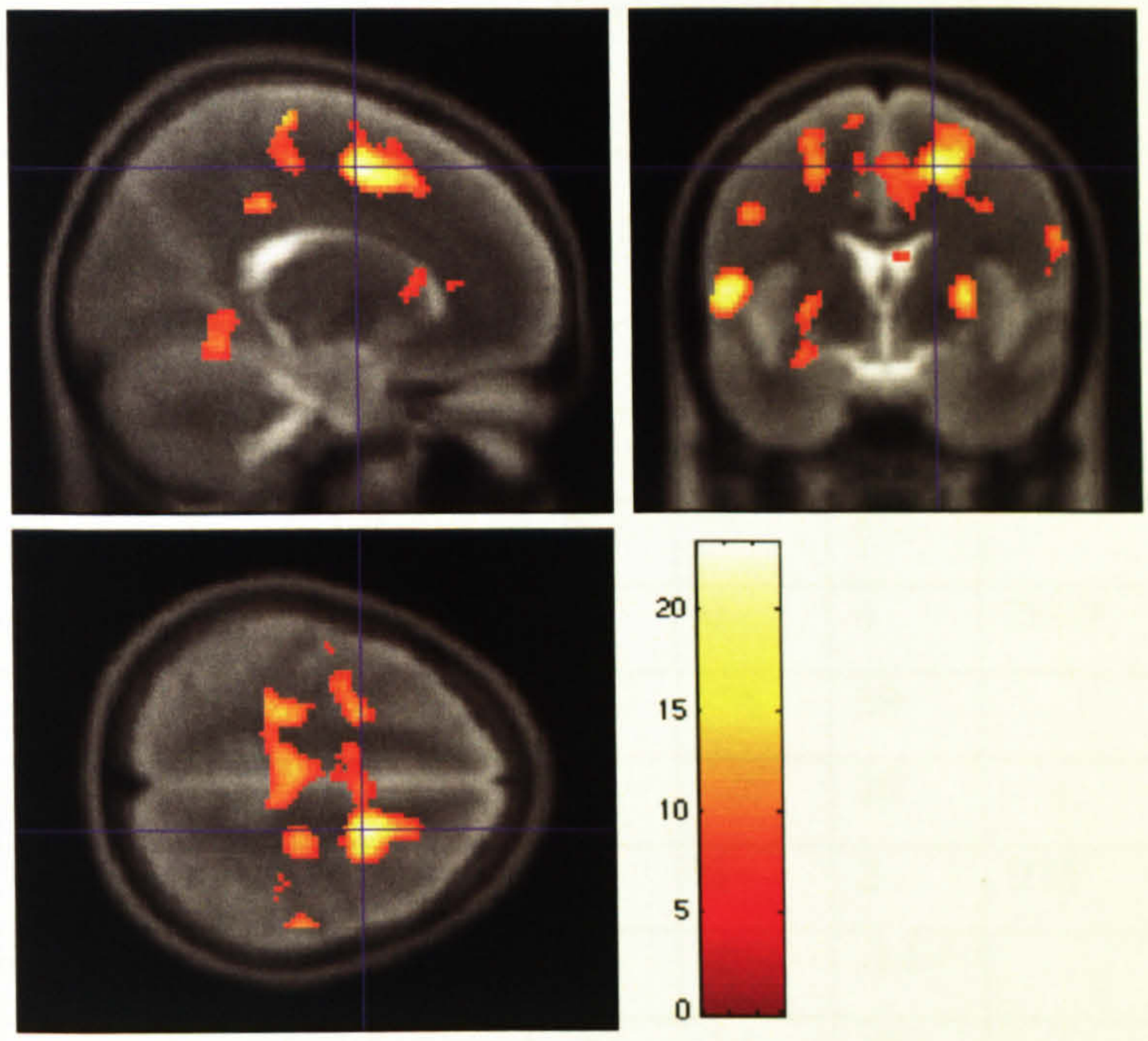

Figure 3. Cortical areas for the young group showing significant differences between correct and error picture-naming ($p < 0.001$ uncorrected).

Region	L/R	MNI coordinates			Voxels (k)	Z
		x	y	z		
Correct > Error						
Anterior Cingulate Gyrus		6	-20	66	7748	6.44
Precentral gyrus	R	62	6	26		5.77
Superior temporal gyrus	R	28	-2	6		5.63
Superior temporal gyrus	L	-58	0	4	2049	5.65
Supermarginal gyrus	L	-56	-38	38		5.36
Supermarginal gyrus	L	-54	-26	20		4.65
Inferior frontal cortex	L	-26	6	2	935	5.09
Superior temporal gyrus	L	30	4	-14		4.61
Sylvian fissure	L	-20	18	-8		4.28
Middle frontal gyrus	R	40	32	32	563	5.41
Precentral gyrus	R	32	24	42		3.36
Middle frontal gyrus	L	-38	34	24	323	4.66
		-38	32	32		4.44
Postcentral gyrus	L	-40	-28	60	250	4.03
Precentral gyrus	L	-24	-4	50		3.84
Inferior parietal lobule	L	-36	-20	-48		3.69
Inferior temporal gyrus	R	22	-54	-6	191	4.30
Inferior temporal gyrus	L	-10	-52	-18	191	4.10
Error > Correct						
Anterior cingulate gyrus		10	12	62	53	3.87
Anterior cingulate gyrus	R	6	18	56		3.30
Anterior cingulate gyrus	L	-2	14	58		3.28
Middle frontal gyrus	L	-52	14	34	46	3.79
Inferior frontal gyrus	L	-52	34	4	14	3.29

Table 5. Cortical areas for the young group showing significant differences between correct and error picture-naming ($p < 0.001$ uncorrected).

6.5 Introduction of Elderly Subjects

6.5.1 Why?

Behavioural studies that have tested the performance of young and elderly subjects on confrontational naming tasks have found evidence of quantitative differences between the two population groups. Older subjects showed reduced accuracy and displayed longer response latencies than their younger counterparts (Tsang and Lee, 2003; Hogson and Ellis, 1998). Functional imaging studies have also reported significant differences in neural activation patterns between healthy young and elderly subjects during cognitive performance. In language tasks specifically, deviations noted included primary language areas such as Broca's and Wernicke's areas and the anterior cingulate cortex and lateral prefrontal cortex, identified as being involved in response monitoring (Fridriksson et al, 2006; Sharp et al, 2006; Persson et al, 2004). Within these studies both increases and decreases in measured activity levels have been reported in key areas, supporting the conclusion that age can affect cortical activation patterns. When this evidence is considered alongside likely age-related neurovascular changes (discussed in Chapter 2) it seems sensible to recruit healthy elderly subjects for a valid comparison of forced naming errors with stroke induced aphasic errors.

6.5.2 Elderly forced error pilot study

A separate pilot study was deemed necessary for the inclusion of elderly participants in an fMRI experiment of forced errors. Considering the evidence that older subjects display longer reaction times on picture naming tasks and that previous speeded naming studies have largely used young students as participants, there was uncertainty as to whether the 600 msec deadline used in the main pilot would be appropriate when applied to older subjects, or if it would just be too short to allow for any kind of response to be produced, errorful or not. For the purpose of this pilot study four different deadlines were trialled to determine which could induce a reasonable level of paraphasias in this population without making the speeded naming task too difficult to perform.

6.5.3 Subjects

There were 7 subjects tested in this study, aged 62-75, with a mean age of 69.57 years (SD = 5.41). They were in good health with no known history of neurological disease or injury. All subjects were predominantly right-handed, as tested by the Edinburgh Handedness Inventory (Oldfield, 1971). Subjects were recruited from an elderly control group who had previously taken part in a study of apraxia and all gave their informed consent. Ethical approval was provided by the University of Nottingham Psychology Department ethics committee.

6.5.4 Experimental procedure

The experiment consisted of four blocks of fifteen naming trials with the response deadline set to 600 msec, 800 msec, 1000 msec and 1200 msec. These times denote the length of the stimulus presentation in each block and the time at which the buzzer sounded. As with the original forced error pilot study the elderly participants were instructed to try and beat the buzzer with their response and to concentrate on the speed of their response rather than the accuracy.

6.5.5 Results

The numbers of errors made by each elderly participant in each of the four experimental conditions are presented in Table 6. The results show that on average the error rate for the different naming deadlines were similar, with only one participant giving an error-free performance for one block. In total only one error was an error of omission, all others were overtly produced error responses. It was surprising to note that the condition with the longest allowable response time was also the most errorful. Although, it is acknowledged that due to the relatively small amount of trials in each block any differences can be exaggerated.

	EP1	EP2	EP3	EP4	EP5	EP6	EP7	Mean (SD)	%
600	5	2	3	2	4	2	3	3 (1.15)	26.58%
800	5	3	2	1	2	4	2	2.71 (1.38)	24.05%
1000	3	3	3	0	2	1	3	2.14 (1.21)	18.99%
1200	6	5	2	2	1	6	2	3.43 (2.15)	30.38%

Table 6. Number of errors made by elderly participants in the four speeded naming conditions with 600, 800, 1000 and 1200 msec naming deadlines. The numbers of errors in each condition are out of 15 naming trials. Also shown are the mean and standard deviations of errors and the percentage of the error total that was produced in each condition. EP stands for elderly participant.

It seems that any restriction on reaction times imposed in this study was enough to enforce errors in this group of seven elderly participants. However, the purpose of this second pilot of the forced error paradigm was to ensure that a deadline of 600 msec, as proposed in the literature and successfully trialled in the main pilot study, would still be valid for use with participants who provide a closer age match to a typical stroke patient. These results suggest that despite evidence of an increase in reaction times in the elderly population such a time limit does not eliminate overtly produced naming responses in an elderly group of participants. Only one omission occurred in the study, indicating that participants felt able to attempt to beat the buzzer as per the task instructions. Therefore, it was decided to keep the same 600 msec naming deadline in line with the pilot and fMRI forced error study of healthy young participants reported above.

6.6 Elderly Forced Error fMRI Study

6.6.1 Subjects

The 7 subjects who partook in the elderly forced error pilot study also participated in the scanning study. It was thought that due to the small amount of speeded naming trials the subjects had already completed and the amount of time that had passed between the elderly pilot and subsequent fMRI study, there would not be any significant practise effects to confound the results.

Ethical approval for this fMRI study was given by the University of Nottingham Psychology Department ethics committee and informed consent was obtained from all participants.

6.6.2 Activation paradigm

The activation paradigm used was identical to that of the previous fMRI study of forced errors using young participants. The elderly participants viewed eighty of Snodgrass and Vanderwart's (1980) black and white line drawings in a randomised order. The stimuli were presented at the rate of one every 15 secs and each was shown for a duration of 600 msec. The removal of the picture at 600 msec into the trial coincided with a buzzer sound that represented the task response deadline. Participants were instructed to name the stimulus item aloud as quickly as possible with the aim of "beating the buzzer". They were advised to concentrate on the speed of their response rather than the accuracy of their naming. During the 14.4 sec inter-stimulus interval (ISI) a fixation cross was shown in the centre of the screen to draw attention to the position of the subsequent stimulus picture. The experiment, consisting of eighty naming trials, lasted for twenty minutes.

The software Presentation (http://neurobs.com) was programmed to control the execution of the experimental paradigm. Within the scanner stimuli were viewed using Avotech "Silent Vision" goggles (http://avotech.org) mounted onto the head coil, the buzzer sound was delivered through headphones worn by the participant and the overt naming responses were picked up by the bespoke microphone setup used in the previous fMRI studies within this project and digitally recorded onto a computer as a .wav file.

6.6.3 Scanning parameters

The scanning protocol was also identical to that already reported for the young forced error fMRI study. A Philips 3T Achieva fMRI scanner with an 8-channel SENSE head coil was used to acquire T2* weighted EPI images of the elderly participants during completion of the experimental paradigm. The field of view included posterior frontal, temporal and anterior parietal cortex and throughout the study 454 image volumes, consisting of 18 coronal 3 x 3 x 6mm slices, were collected for analysis. The volume repetition time (TR) was 2.00s with an echo time (TE) of 35ms and the sense factor was 2.

6.6.4 Data analysis

The processing of recorded spoken task responses and the pre-processing and statistical analysis methods used for the functional imaging data was identical to those used in the young forced error study.

The data from EP4 was excluded from the analysis due to half of EP4's responses being missing from the sound recording. It was not clear if EP4 stopped responding or if the microphone had shifted during the scan session to prevent the speech from being detected. The remaining six datasets were entered into the analysis.

The digital recordings containing participants' naming responses were run through the noise cancellation tool (Cusack et al, 2005) to reduce the impact of the scanner noise on the integrity of the recorded speech. Responses were transcribed and categorised according to the pre-existing classification system used in several previous studies of aphasia and throughout this research project. Exact timings for the onset of naming responses (in minutes, seconds and milliseconds) was extracted from the sound file for the purpose of building an accurate model of the underlying cortical activity during analysis of the fMRI signal. Naming responses across all response categories were grouped more generally into corrects and errors, representing the two event types to be statistically compared in the fMRI data analysis.

All pre-processing and statistical analysis of the fMRI data took place within SPM5 (http://www.fil.ion.ucl.ac.uk/spm/). Images collected from the elderly forced error participants were corrected for slice timing differences in the acquisition of each image volume and for any head motion that may have occurred within the scanner. Images were then spatially normalised to MNI space and smoothed using a gaussian kernel with a FWHM of 8mm. The data was finally passed through a temporal filter with a 30s cut-off point, removing uninteresting low frequency variability before the statistical analysis is performed.

The two pre-defined response groupings of corrects and errors were entered into the model as separate conditions, using the extracted speech onset times to describe the beginning of a 1s task related stimulation. Images from the individual elderly participants were combined together to form a fixed effects group analysis. Activation maps were

generated for each condition ($p < 0.05$, using family-wise error (FWE) correction for multiple comparisons). All areas activated in both the correct or error conditions were combined into a statistical map produced using a threshold of $p < 0.001$ uncorrected. The resulting image was saved as a masking image for use in the small volume correction to reduce the number of voxels in the search space to those showing task related signal change and therefore reduce the number of statistical tests performed.

Statistically significant differences in activations between the correct and error response trials were investigated by directly comparing the conditions with the contrasts error > correct and correct > error ($p < 0.001$, uncorrected, masked by an inclusive mask of (error and correct) conditions at $p < 0.001$, uncorrected).

6.6.5 Results

6.6.5.1 Elderly naming performance

The elderly forced error group achieved an overall mean performance level of 58.67 correct (SD = 5.35) and 21 error naming responses (SD = 4.90). Table 7 shows that elderly participants were quicker to produce correct names, although only one subject managed to produce a name before the 600 msec deadline in either condition.

	Correct Responses			Error Responses		
	Mean (SD)	**Range**	**N**	**Mean (SD)**	**Range**	**N**
EP1	0.80 (0.13)	0.56-1.13	52	0.79 (0.13)	0.62-1.08	26
EP2	1.24 (0.42)	0.67-3.05	62	1.98 (0.86)	1.20-3.59	18
EP3	0.87 (0.21)	0.61-1.88	66	1.12 (0.57)	0.65-2.23	14
EP5	0.87 (0.26)	0.62-1.97	59	1.08 (0.43)	0.64-2.20	21
EP6	1.09 (0.28)	0.72-1.85	60	1.36 (0.34)	0.67-1.99	20
EP7	0.93 (0.31)	0.60-2.38	53	1.18 (0.56)	0.63-3.17	27

Table 7. Speech onset times for correct and error naming responses measured in seconds and the number of corrects and errors produced by each elderly participant.

Table 8 shows that errors made by this participant group were mostly semantic in nature (35.2% share with a further 12.8% of semantic/visual errors). The elderly group made a high number of indistinguishable errors, with no responses and other errors also featuring. Very few non-word and formally related errors were produced.

Error Type	EP1	EP2	EP3	EP5	EP6	EP7	%
Semantic	7	9	3	9	7	9	35.2
Semantic/Visual	3	1	3	3	3	3	12.8
Formal	0	0	0	0	0	0	0
Mixed	1	0	0	0	1	0	1.6
Unrelated	0	1	1	1	1	3	5.6
Unrelated/Visual	1	0	0	1	0	1	2.4
Non-word	0	0	1	0	0	0	0.8
No response	6	4	4	0	0	0	11.2
Indistinguishable	8	1	2	3	2	9	20
Other	0	2	0	3	6	2	10.4

Table 8. Breakdown of the number of different error types made by the elderly participants, with the percentage of error trials represented by each error type.

6.6.5.2 Elderly fMRI results

Several areas showed large activations for the correct naming condition in the healthy elderly forced error group (see Figure 4). These included precentral gyrus and inferior frontal gyrus bilaterally, with greater activation in the left hemisphere, and the anterior cingulate gyrus. The inferior and superior temporal gyri and cerebellum also showed bilateral activation (see Table 9).

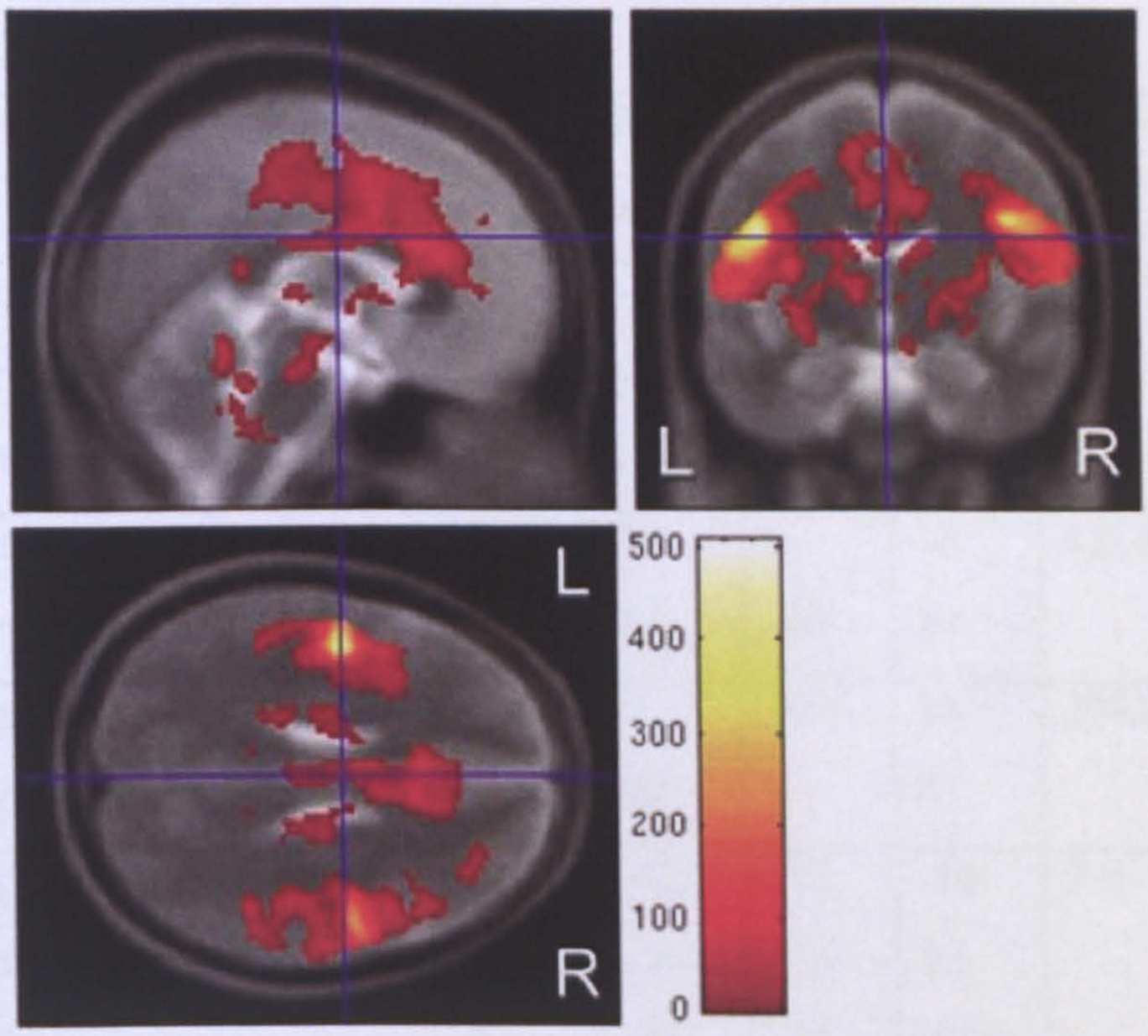

Figure 4. Cortical areas for the healthy elderly group showing significant activation during correct picture-naming ($p < 0.05$ FWE corrected).

Region	L/R	MNI coordinates			Voxels (k)	Z
		x	y	Z		
Precentral Gyrus (BA 6)	L	-48	-10	30	6342	27.4
Precentral Gyrus (BA 4)	L	-54	-6	20		21.5
Inferior Frontal Gyrus (BA 44)	L	-54	10	6		17.8
Inferior Frontal Gyrus (BA 44)	R	64	0	18	4214	28.7
Precentral Gyrus (BA 6)	R	44	-10	34		22.0
Anterior Cingulate Gyrus (BA 32)		2	30	18	2982	13.4
Anterior Cingulate Gyrus (BA 24)		10	8	6		12.2
Inferior Temporal Gyrus (BA 37)	L	-42	-52	-14	748	15.7
Superior Temporal Gyrus (BA 22)	L	-58	-38	12		14.6
Cerebellum	L	-36	-48	-24		13.0
Cerebellum	L	-16	-52	-16		12.9
Inferior Temporal Gyrus (BA 37)	R	34	-50	-20	702	15.2
Cerebellum	R	16	-50	-6		10.5
Thalamus	R	18	-12	-8	437	9.5
Thalamus	R	8	-14	4	123	7.7
Superior Temporal Gyrus (BA 22)	R	56	-38	10	37	8.3
Anterior Insula (BA 13)	R	34	-10	10	16	6.7
Heschl's Gyrus (BA 41)	R	40	-28	12	14	7.0

Table 9. Summary of cortical activations associated with correct naming by the healthy elderly group ($p < 0.05$ FWE corrected).

Figure 5 presents the areas that showed significant differences between the two naming response conditions. Table 10 shows that correct naming responses selectively activated the cerebellum and left inferior temporal gyrus. The production of error responses was associated with activation of anterior cingulate gyrus and insula and the inferior frontal gyrus bilaterally, with greater activation on the left side.

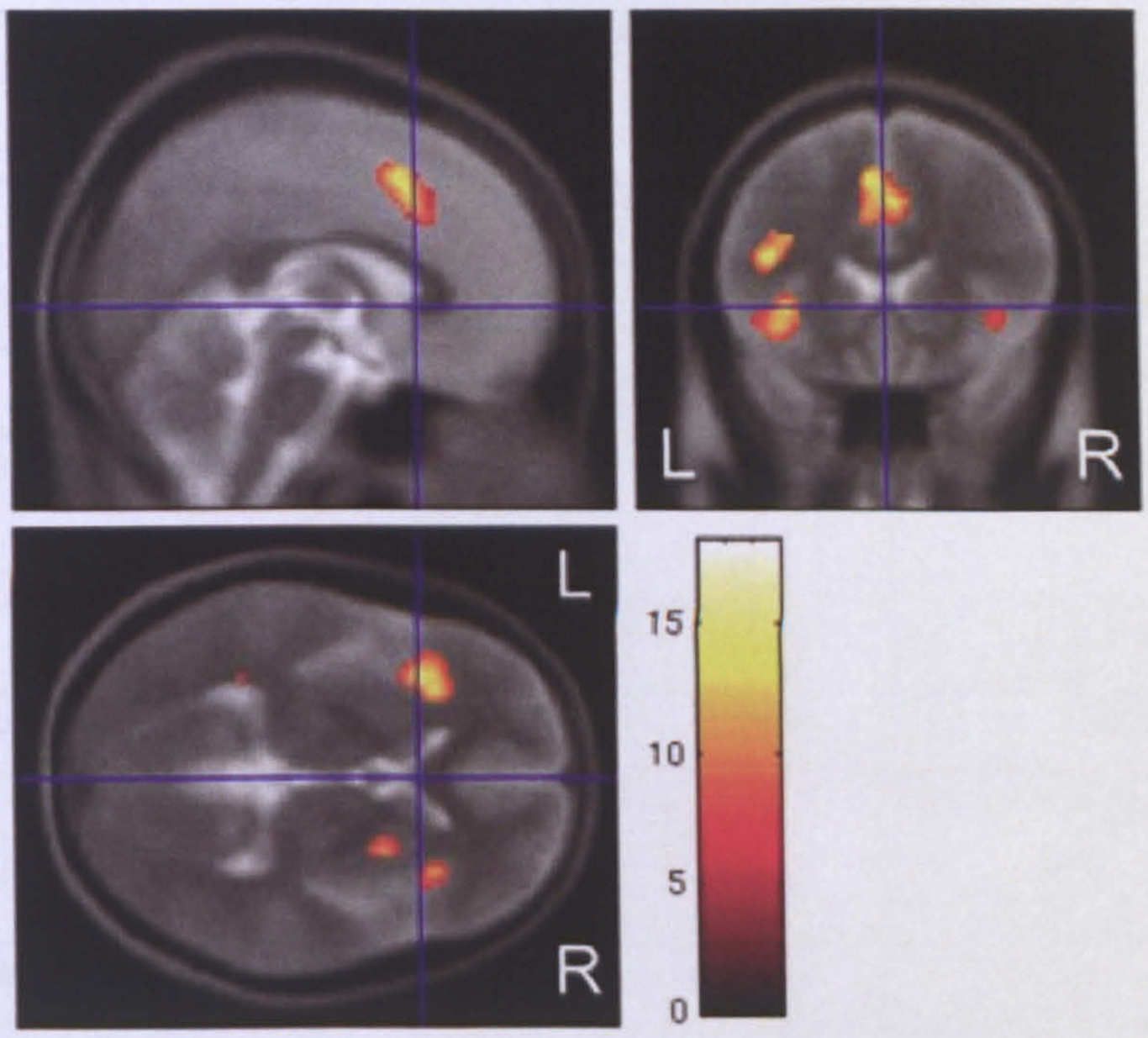

Figure 5. Cortical areas for the healthy elderly group showing significant differences between correct and error picture naming ($p < 0.001$ uncorrected).

Region	L/R	MNI coordinates			Voxels (k)	Z	P
		x	y	z			
Correct > Error							
Cerebellum	R	18	-48	-24	146	9.5	<0.001
Inferior Temporal Gyrus (BA 37)	L	-48	-44	-6	93	9.3	<0.001
Putanem	R	26	8	-2	55	9.0	<0.001
Precentral Gyrus (BA 6)	R	48	-6	30	17	6.5	0.008
Middle Frontal Gyrus (BA 46)	R	50	44	6	10	6.3	0.012
Error > Correct							
Anterior Cingulate Gyrus (BA 32)		-4	16	48	485	10.5	<0.001
Insula (BA 13)	L	-36	22	0	402	10.2	<0.001
Inferior Frontal Gyrus (BA 45)	L	-42	22	18	348	9.7	<0.001
Inferior Frontal Gyrus (BA 47)	R	34	26	0	86	9.1	<0.001
Medial Frontal Gyrus (BA 9)	R	34	12	36	29	7.4	0.003

Table 10. Summary of cortical activations reaching significance for the contrasts (correct > error) and (error > correct) by the healthy elderly group ($p < 0.001$ uncorrected). MNI co-ordinates, cluster size, maximum z score and p value adjusted using small volume correction are presented.

6.7 Comparison of Young and Elderly

6.7.1 Naming performance

A comparison of mean response times for correct and error naming conditions between the two groups of participants shows that, as expected, in general the young participants were faster at producing responses than the elderly group (see Figure 6). The elderly group showed a more consistent difference in response times between the correct and error conditions.

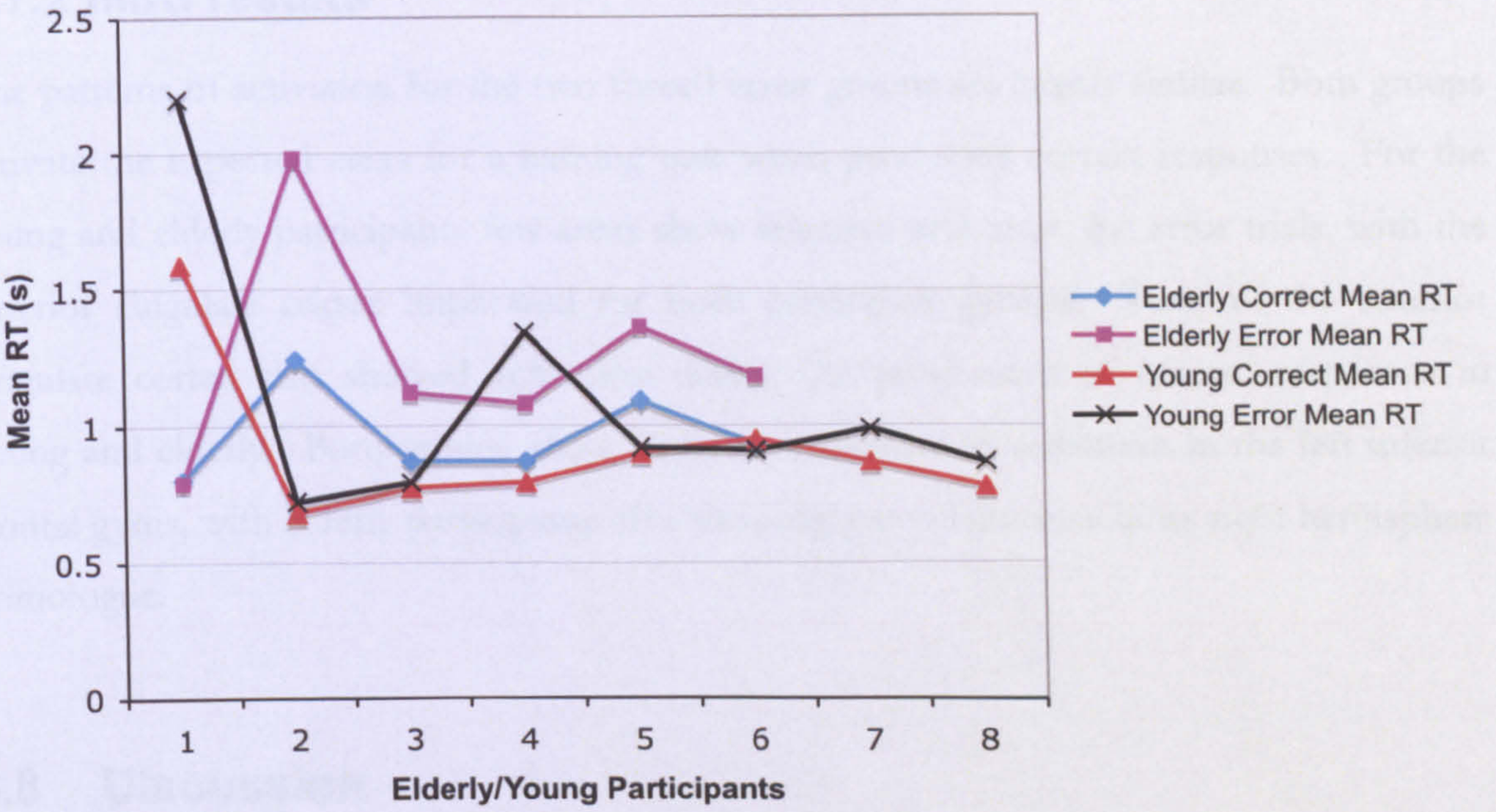

Figure 6. Line chart showing the mean response times (s) for correct and error names in the elderly and young forced error groups.

Figure 7 shows that both young and elderly participants mostly produced semantic errors to the naming-to-deadline task. The young participants inhibited their response more often when they were unsure of the correct target name.

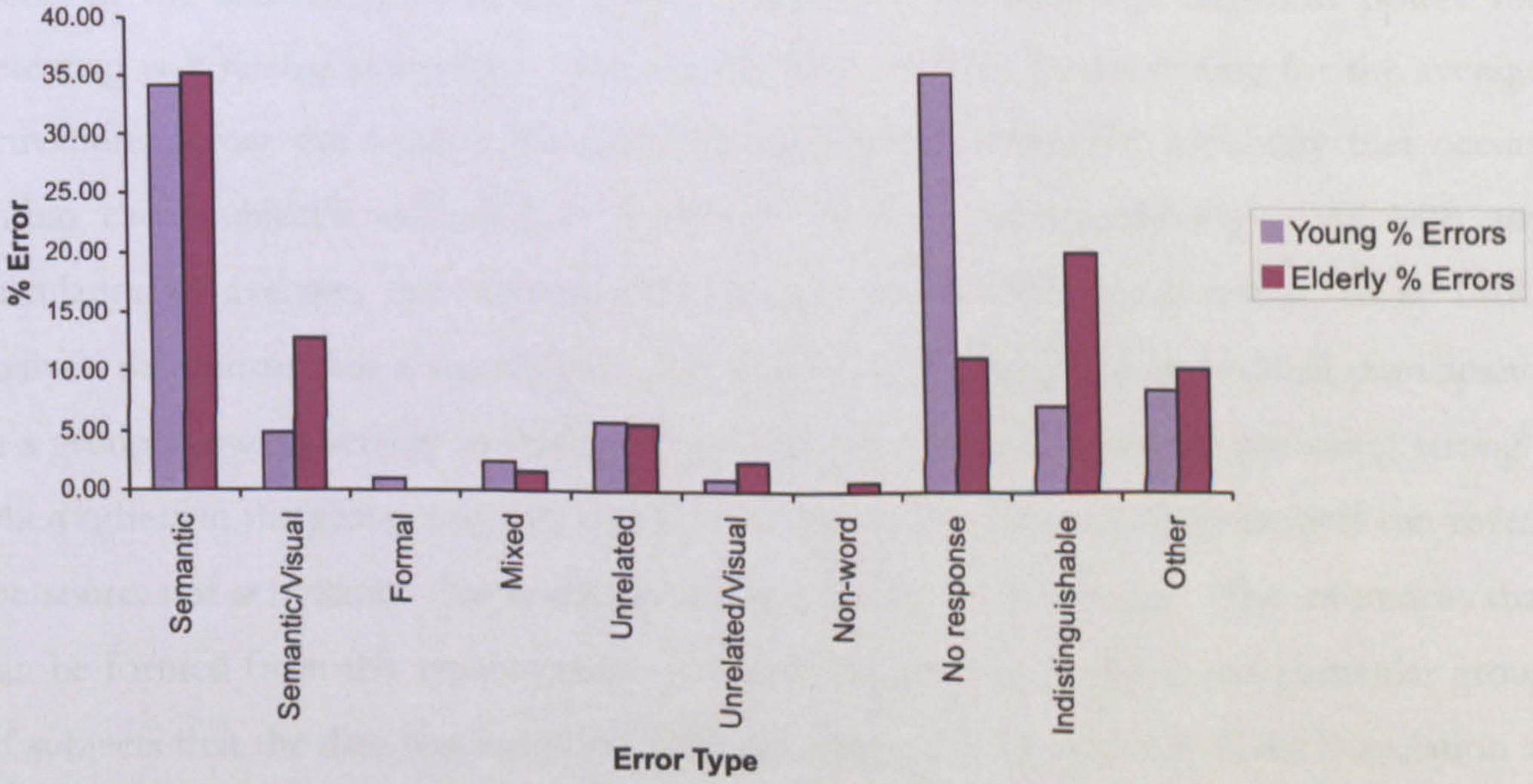

Figure 7. Bar chart showing the percentage of different error types produced by the elderly and young forced error groups.

6.7.2 fMRI results

The patterns of activation for the two forced error groups are largely similar. Both groups activate the expected areas for a naming task when producing correct responses. For the young and elderly participants few areas show selective activation for error trials, with the anterior cingulate cortex implicated for both participant groups. Parts of the anterior cingulate cortex also showed activation during the production of correct responses in young and elderly. Both groups show a selective increase in activation in the left inferior frontal gyrus, with elderly participants also showing a small increase in its right hemisphere homologue.

6.8 Discussion

To the author's knowledge, these studies represent the first time a naming deadline paradigm has been used in a functional imaging experiment.

Before attempting to interpret the results it is important to acknowledge the constraints that the statistical method used places on the kind of inferences that can be made. In this study of forced errors the two participant groups, young and elderly, were analysed separately in two fixed effects analyses. Grouping participant data together in this way increases the sensitivity to neural signal changes by boosting the statistical power for detecting task related activations. In a fixed effects analysis we are testing for the average activations across the subject group, taking into consideration the variability that occurs within each subject's individual scan session (scan to scan variability). As with any calculation of averages any outlying data points can skew the overall result. In an fMRI analysis this means that a significantly activated brain area could arise from all participants in a group showing activity in this area, or just a proportion of subjects activating strongly when others in the group may not activate this area at all. Single subject analysis can reveal the sources of activations that reach significance in the average image. The inferences that can be formed from this type of analysis, therefore, can only apply to the particular group of subjects that the data was extracted from and cannot be generalised to the population as a whole that the subjects were drawn from.

To allow for conclusions to be applied to the population whose characteristics are being studied, i.e. people who have suffered a stroke, the fMRI data needs to undergo a random

effects analysis. This type of analysis aims to identify the areas that activate similarly in all participants, or scan sessions, within a group. Instead of assuming that the different contributions from each subject are "fixed", activations are thought of as random variables and the between session/subject variability in activation is included in the statistical model. Therefore, in a random effects analysis strong activations that only show in a small proportion of the subjects in a group will be recognised as a variation from the group average and will not reach significance. Modelling the variability of activations between subjects means that it is now valid to make inferences from the results about the whole population that the subject group was drawn from (Friston et al, 1999).

Although a random effects analysis may seem like the obvious choice for assessing group data, its use is not always appropriate. In order to apply study results to the population of interest it is necessary to accurately estimate the between-subject variance. The variability that exists between different subjects will obviously be greater than the scan-to-scan variability considered in a fixed effects analysis, and also in the first stage of a random effects analysis. The implication of this for random effects models is that the more subjects studied, the more robust the estimate. To perform a fixed effects group analysis the data from each subject is simply grouped together in a model, summing the number of data points and degrees of freedom from each single subject analysis and greatly improving statistical power. To enter fMRI group data into a random effects analysis each subject's individual results are summarised with a single image which are then entered into the second stage model to assess the between subject variability. The number of independent observations submitted into this model has now been reduced to correspond with the number of subjects in the group. To illustrate this power difference, in this fMRI study of forced errors there were 8 participants in the young group and 6 in the elderly group, this corresponds to a total of 3632 scans for the young group and 2724 for the elderly that were entered into the fixed effects model.

In a study of statistical power in functional neuroimaging data Van Horn et al (1998) performed power calculations on images from ever decreasing subsets taken from an initial group of 40 participants. They concluded that for the cortical areas involved in their activation task to show reliable cognitive activation a group sample size approaching 20 participants was required. It has also been reported that groups of less than about 10 subjects are generally considered to be suboptimal for entering into a random effects

analysis (Smith, 2001). Since the sample sizes achieved in this study were relatively small for a group study it was considered that there simply would not be sufficient power to detect cortical activation in response to the task paradigm using random effects statistics. On this occasion it was deemed necessary to forgo the validity and generalisation afforded by a random effects analysis in favour of the ability to detect task related signal changes at all. Therefore, these results must be interpreted within the context of the specific participants whose images were acquired.

The results from this study show that, as predicted, healthy participants who are artificially induced into increasing their error rate show increases in activation in areas that have been associated with the monitoring and control of errors in unimpaired subjects (e.g. MacDonald et al, 2000). These areas were also activated in some of the aphasic patients specifically during the production of errors (AP3 and AP4) and in one patient throughout the naming task (AP5). This suggests that the naming-to-deadline paradigm can be successfully used as a baseline for control groups in studies of aphasic naming to control for the normal error processing responses and make abnormal aphasic responses in these and other brain areas easier to interpret.

Some limited recruitment of right hemisphere language homologues in error trials was found, though only in the elderly group where comparatively small areas of the inferior and medial frontal gyri showed a signal increase. Correct naming activated the normal distributed, bilateral language network. This finding suggests that placing a high load on normal language processing in this way does not simulate activations seen in aphasic patients after damage as hypothesised. If this were the case predicted loci of activity, such as right inferior frontal and middle and superior temporal areas, should have been activated in both young and elderly participant groups. As this did not occur the results from the forced error studies go against implications made by the connectionist models of language that view normal and abnormal language on the same continuum. Results instead suggest that there is a qualitative difference between error brain patterns of normals and patients rather than simply a quantitative variation.

The implication of this evidence for the overall aims of this research project is that to study the neural basis of the language system in the presence of stroke damage, and how recovery of function can occur, it is not useful to draw conclusions from studies of healthy

participants unless a more realistic method of artificially impairing healthy performance can be found. In terms of recovery from aphasia, the forced error studies suggest that the stroke damage causing the deficit has resulted in a real change in the cortical organisation of language processing, rather than just a reduction in efficiency of the system.

7 General Discussion

7.1 Introduction

This thesis represents a progression of research from the conception and development of a novel fMRI paradigm through behavioural and functional imaging pilot work (Chapters 3 and 4) to allow for the within-subject comparison between correct and error confrontational naming performance in aphasic patients (Chapter 5) and subsequently with time stressed unimpaired participants (Chapter 6). It was thought that these functional imaging studies would support previous research that implicated an association between increased activity in right hemispheric contralesional areas and the production of aphasic errors. Conversely, it was expected that successful picture naming would coincide with increased activation in left hemisphere perilesional sites. Unimpaired participants, who performed the same task but with a predefined response deadline, were expected to recreate patient activation patterns. This hypothesis was based on the theoretical notion that aphasic output represents overall weaknesses in the language processing network that can also occur when the normal system is overloaded.

The overall aims of this research were to successfully acquire data from aphasic stroke patients using an event-related, continuous scanning, overt naming paradigm and to provide a significant contribution to the current debate in the literature concerning the role of the right hemisphere in functional recovery from aphasia. Although the experimental hypotheses were not fully supported by the data from either patients or healthy forced error participants, it is judged that the results as presented here fulfil these aims and provide useful suggestions for the direction future research should take and has implications for therapeutic practice.

7.2 Evaluating the Paradigm

A simple picture-naming task was chosen for the purpose of this project as people with aphasia commonly complain of word-finding difficulties in everyday conversations and an aphasic deficit is often revealed in assessments of confrontational naming through the

production of errors and omissions (Howard and Gatehouse, 2006). It is also a relatively simple task to set up and control within a functional imaging study and naming responses can easily be recognised and categorised as being correct or an error. Formulae for categorising naming responses are well established in the literature (Dell et al, 1997; Caramazza et al, 2000; Capitani & Laiacona, 2004) and there is good agreement of the neural correlates associated with normal naming performance (Indefrey & Levelt, 2004; Wise, 2003; Hillis, 2007).

The picture-naming task used successfully evoked a significant number of error responses in the aphasic patients who participated in the main fMRI study. This allowed for direct contrasts to be made between brain areas associated with correct and error processing.

This experimental procedure represented an improvement on previous studies that did not collect spoken responses from participants (e.g. Ellis et al, 2006; Perani et al, 2003). Requesting for covert responses to language tasks has been found to be an unacceptable compromise to solve the problem of speech-induced motion artifact (Huang et al, 2001). The separation of correct and error trials in data analysis may also be a more valid and revealing way of investigating recovery of functioning, rather than lumping all response types together or excluding trials containing error processing. There are now a large number of studies that have looked at recovered language processing and so to make advances in this field of research more analytical experimental methods should be explored.

After the completion of this research two other attempts at comparing brain activity from correct and error responses were made, suggesting that other researchers also consider this to be an advantageous route for research into recovery from aphasia. Meinzer et al (2006) reported a case study of one patient with chronic aphasia who was tested pre and post-training with an overt picture-naming task where correct naming, semantic paraphasias and neologisms were separated in analysis. 26 instances of each response type were chosen post-hoc to balance the statistical power achieved by each condition. However, this study used a sparse sampling procedure that necessarily means some information about the haemodynamic response function from a portion of every trial will be missing. This may prove costly when such a paradigm is applied across a group of aphasic participants, where speech onset times can vary greatly.

In their review of the use of fMRI in aphasia research Crosson et al (2007) reported that they had made a failed attempt to compare correct and error trials in a study using continuous sampling. They found that they did not collect enough instances of each response type to enter into statistical analysis. As far as it is known the patient study presented in this thesis therefore represents the first successful attempt at providing continuous fMRI datasets contrasting different linguistic response types.

7.3 The Left Versus Right Debate

Evidence supporting each side of the debate was presented in Chapter 1. Results from the patient study broadly support the studies that suggest right hemisphere language areas can make a useful contribution to the return of language function in aphasia (Basso et al, 1989; Gold et al, 2000; Musso et al, 1999). This finding does not discount the idea that left hemisphere perilesional areas are also important for recovery and, indeed, the one patient who achieved the best naming performance in the study did show an association between increased perilesional activation and correct responding. This supports the idea of reactivation of the normal language processing network as far as is possible after stroke damage being the most successful route to recovery. This particular patient's data also included an increase of activity in right frontal areas, homologous to the lesion site, when errors were being produced. This means that the results from all five of the aphasic patients cannot completely refute the argument for the release of right hemisphere areas being incidental or even detrimental to language performance (Price and Crinion, 2005; Naeser et al, 2005).

Given the mixture of conclusions that can be reached from viewing the data from the aphasic patients individually, the results from this research support the position put forward by Crosson et al (2007); that theories of recovery must leave behind the simplistic left versus right battle and focus more on the circumstances under which perilesional areas can be reactivated to promote recovery and those where the reorganisation of language processing to right hemisphere homologues can be effective. Uncovering this for individual aphasic patients could be used to advantage in rehabilitation, to allow therapists to tailor their interventions according to the most advantageous recovery strategy for the patient's damaged cortical language network.

The study by Meinzer et al (2006) found that differences between correct and error naming mainly involved right inferior frontal areas. Specifically, increased activation in the right inferior frontal gyrus was found to be related to better naming performance. Some results from the patient study (AP1) tentatively support this finding, but only in one single subject. The patient study results, however, generally agree with Meinzer et al's conclusion of useful processing occurring in right hemisphere language homologues.

7.4 Is It Useful to Study Forced Errors?

Some recent psycholinguistic models of language have considered normal and abnormal processing to be bound within the same architecture, with linguistic deficits representing inefficiency in the normal system rather than a significant departure from unimpaired processing. If this is true then, logically speaking, placing stresses on the normal system should bring about task performance similar to that seen in aphasia. Forced error participants, however, differed from the aphasic patient performance in the levels of formal/non-word type errors made. Very few of these errors were produced by the forced error groups compared to the patients, with errors induced under temporal stress tending to be more semantic in nature. Errors produced by the naming-to-deadline paradigm are likely to represent the processing stage that is constrained by this task design and it is logical that this will result in similar error types being produced by different participants. Considering the behavioural and imaging results from this study it is hard to see how this task could induce the variable error patterns and response times seen in patients. However, the activation of areas often associated with normal error processing and performance monitoring (e.g. anterior cingulate cortex) suggest that this task may make an improved baseline measure when comparing unimpaired control groups with impaired aphasic task performance.

Sharp et al (2004) deliberately degraded the performance of their normal participants by varying the task difficulty in their PET study of the semantic system. This was done to attempt to reduce the confound of making comparisons between an impaired patient group and the near perfect task performance levels of a control group. Sharp et al suggest that investigating the effects of varying performance levels in normal subjects can aid interpretation of patient data.

In this way, inducing picture-naming errors in unimpaired participants could provide a more valid baseline for making comparisons with aphasic performance levels. A forced error control group can show cortical activations associated with the normal error monitoring and performance control mechanisms so that differences between patient and control activations are more likely to be a result of abnormal processing than simply comparing against ceiling level performances.

7.5 Future Directions

This research shows that successful continuous sampling of aphasic patients and subsequent analysis contrasting correct and error naming performances is possible. However, this study did fall foul of motion artifact to a degree and so further research into reducing the impact of speech-related motion on fMRI data would be very beneficial to this research area.

In terms of the specific study design used in this piece of research, a criticism of the study would be that responses were only separated into generic correct and error categories. This was due to the simple fact that there were not enough instances of each different error type to allow for further separation, e.g. into categories of semantic, formal etc. If this were possible then it would be valuable in the future to further the work done by Meinzer et al and investigate the different processing routes for different types of error responses. Results from such a study could suggest targeted interventions, using technologies such as transcranial magnetic stimulation, matched to a patient's specific error profile in an attempt to alter the faulty processing route to a more fruitful one.

Within this project five aphasic patients were studied individually to test the experimental hypothesis. Increasing the number of subjects may allow for identification of patterns of neural and behavioural characteristics shared by patients that could be associated with a particular type of cortical reorganisation of language processing. For example, lesions in certain areas may increase the likelihood of useful contralateral reorganisation of function. An extension to such a study would be to recruit groups of stroke patients based on lesion location and/or aphasic symptoms to investigate whether these variables affect the neural recovery strategy employed by the brain. Data from these experiments would seek to answer the question of under which circumstances perilesional or contralateral

reorganisation can successfully contribute to recovery (Crosson et al, 2007) and could suggest what rehabilitative strategy would be most effective for patients with certain neurological and behavioural profiles. These studies, however, would be practically challenging to conduct due to the relative difficulty of recruiting and imaging a large number of aphasic stroke patients within a reasonable time frame. Also, the inhomogeneous nature of the aphasic population would make it difficult to group patients together on the basis of an isolated variable, whilst controlling for the effect of others.

The work done within this research project on inducing errors from healthy, unimpaired participants suggests that this method may be a useful high-level baseline measure to employ in future studies of aphasic language errors. Adding in the additional control for processes involved in normal error monitoring and processing could further help to delineate areas specifically contributing to disordered language performance. For this reason a future study directly contrasting forced errors and aphasic errors in a single analysis would be very interesting to conduct.

The research presented here highlights the potential for the use of fMRI in therapeutic diagnosis. If, as here, brain areas associated with better or poorer performance can be delineated rehabilitation can focus on the promotion or inhibition of these areas. This can potentially be done using repetitive treatment with transcranial magnetic stimulation to encourage reorganisation of processing into a more effective activation pattern. Studies trialing this method have reported encouraging results (e.g. Naeser et al, 2005).

7.6 Concluding Statement

The motivation for designing, developing and successfully applying this study paradigm with aphasic patients and unimpaired participants, in whom errors were induced, was to improve on recently reported methodologies, contribute to the debate on recovery from post-stroke aphasia and provide a platform for future studies in this research area. It is believed that these aims have been achieved and that this body of research makes a useful addition to the literature concerning the post-stroke recovery of language.

References

"What is a stroke?". The Stroke Association, May 2007.

Abo, M., Senoo, A., Watanabe, S., Miyano, S., Doseki, K., Sasaki, N., Kobayashi, K., Kilkuchi, Y. and Yonemoto, K. (2004). Language-related brain function during word repetition in post-stroke aphasics. *Neuroreport.* 15:12, 1891-1894.

Abrahams, S., Goldstein, L. H., Simmons, A., Brammer, M. J., Williams, S. C., Giampietro, V. P., Andrew, C. M. and Leigh, P. N. (2003). Functional magnetic resonance imaging of verbal fluency and confrontation naming using compressed image acquisition to permit overt responses. *Human Brain Mapping.* 20, 29-40.

Ansaldo, A. I. and Arguin, M. (2003). The recovery from aphasia depends on both the left and right hemispheres: Three longitudinal case studies on the dynamics of language function after aphasia. *Brain and Language.* 87, 177-178.

Bakas, T., Kroenke, K., Plue, L. D., Perkins, S. M. and Williams, L. S. (2006). Outcomes among family caregivers of aphasic versus nonaphasic stroke survivors. *Rehabilitation Nursing.* 31:1, 33-42.

Bandettini, P. A. and Cox, R. W. (2000). Event-related fMRI contrast when using constant interstimulus interval: Theory and experiment. *Magnetic Resonance in Medicine.* 43, 540-548.

Barch, D. M., Sabb, F. W., Carter, C. S., Braver, T. S., Noll, D. C. and Cohen, J. D. (1998). Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. *NeuroImage.* 10, 642-657.

Barch, D. M., Braver, T. S., Sabb, F. W. and Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience.* 12, 298-309.

Barry, C., Hirsh, K. W., Johnston, R. A. and Williams, C. L. (2001). Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *Journal of Memory and Language.* 44:3, 350-375.

Basso, A., Gardelli, M., Grassi, M. P. and Mariotti, M. (1989). The role of the right-hemisphere in recovery from aphasia – 2 case studies. *Cortex.* 25:4, 555-566.

Best, W., Herbert, R., Hickin, J., Osborne, F. and Howard, D. (2002). Phonological and orthographic facilitation of word-retrieval in aphasia: Immediate and delayed effects. *Aphasiology.* 16:1-2, 151-168.

Birn, R. M., Bandettini, P. A., Cox, R. W., Jesmanowicz, A. and Shaker, R. (1998). Magnetic field changes in the human brain due to swallowing or speaking. *Magnetic Resonance in Medicine.* 40, 55-60.

Birn, R. M., Bandettini, P. A., Cox, R. W. and Shaker, R. (1999). Event-related fMRI of tasks involving brief motion. *Human Brain Mapping.* 7, 106-114.

Birn, R. M., Cox, R. W. and Bandettini, P. A. (2002). Detection versus estimation in event-related fMRI: Choosing the optimal stimulus timing. *NeuroImage.* 15, 252-264.

Birn, R. M., Cox, R. W. and Bandettini, P. A. (2004). Experimental designs and processing strategies for fMRI studies involving overt verbal responses. *NeuroImage.* 23, 1046-1058.

Blackschaffer, R. M. and Osberg, J. S. (1990). Return to work after stroke – Development of a predictive model. *Archives of Physical Medicine and Rehabilitation.* 71:5, 285-290.

Blank, S. C., Bird, H., Turkheimer, F., Wise, R. J. S. (2003). Speech production after stroke: The role of the right pars opercularis. *Annals of Neurology.* 54, 310-320.

Bock, K. and Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). New York: Academic Press.

Boller, F. and Vignolo, L. A. (1966). Latent sensory aphasia in hemisphere damaged patients: An experimental study with the Token Test. *Brain.* 89, 815-830.

Buckner, R. L. (1998). Event-related fMRI and the hemodynamic response. *Human Brain Mapping.* 6:5-6, 373-377.

Capitani, E., Laiacona, M., 2004. A method for studying the evolution of naming error types in the recovery of acute aphasia: a single-patient and single-stimulus approach. *Neuropsychologia.* 42, 613-623.

Caramazza, A., Papagno, C., Ruml, W., 2000. The selective impairment of phonological processing in speech production. *Brain and Language.* 75, 428-450.

Carusone, L. M., Srinivasan, J., Gitelman, D. R., Mesulam, M. M. and Parrish, T. B. (2002). Hemodynamic response changes in cerebrovascular disease: Implications for functional MR imaging. *AJNR: American Journal of Neuroradiology.* 23, 1222-1228.

Christoffels, I. K., Formisano, E. and Schiller, N. O. (2007). Neural correlates of verbal feedback processing: An fMRI study employing overt speech. *Human Brain Mapping.*

Cornelissen, K., Laine, M., Tarkiainen, A., Jarvensivu, T., Martin, N. and Salmelin, R. (2003). Adult brain plasticity elicited by anomia treatment. *Journal of Cognitive Neuroscience.* 15:3, 444-461.

Cramer, S. (2004). Functional imaging in stroke recovery. *Stroke.* 35:11, 2695-2698.

Cornelissen, K., Laine, M., Tarkiainen, A., Jarvensivu, T., Martin, N. and Salmelin, R. (2003). Adult brain plasticity elicited by anomia treatment. *Journal of Cognitive Neuroscience.* 15:3, 444-461.

Crosson, B., McGregor, K., Gopinath, K. S., Conway, T. W., Benjamin, M., Chang, Y-L., Bacon Moore, A., Raymer, A. M., Briggs, R. W., Sherod, M. G., Wierenga, C. E. and White, K. D. (2007). Functional MRI of language in aphasia: A review of the literature and the methodological challenges. *Neuropsychology Review.* 17, 157-177.

Cusack, R., Cumming, N., Bor, D., Norris, D., Lyzenga, J., 2005. Automated post-hoc noise cancellation tool for audio recordings acquired in an MRI scanner. *Human Brain Mapping.* 24, 299-304.

Dale, A. M. and Buckner, R. L. (1997). Selective averaging of rapidly presented individual trials using fMRI. *Human Brain Mapping.* 5, 329-340.

Damasio, H., Grabowski, T. J., Tranel, D., Hichwa, R. D. and Damasio, A. R. (1996). A neural basis for lexical retrieval. *Nature.* 380, 499-505.

DeLeon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., Lee, A. and Hillis, A. E. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain.* 130, 1408-1422.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M. and Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review.* 104:4, 801-838.

Dell, G. S., Martin, N. and Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language.* 56, 490-520.

D'Esposito, M., Zarahn, E., Aguirre, G. K. and Rypma, B. (1999). The effect of normal aging on the coupling of neural activity to the BOLD hemodynamic response. *NeuroImage.* 10, 6-14.

D'Esposito, M., Ballard, D., Zarahn, E. and Aguirre, G. K. (2000). The role of prefrontal cortex in sensory memory and motor preparation: An event-related fMRI study. *NeuroImage.* 11, 4000-408.

D'Esposito, M., Deouell, L. Y. and Gazzaley, A. (2003). Alterations in the BOLD fMRI signal with ageing and disease: A challenge for neuroimaging. *Nature Reviews: Neuroscience.* 4, 863-872.

Dronkers, N. F. (2000). The gratuitous relationship between Broca's aphasia and Broca's area. *Behavioural and Brain Sciences.* 23:1, 30-31.

Eden, G.F., Joseph, J .E., Brown, H.E., Brown, C.P. and Zeffiro, T.A. (1999). Utilising hemodynamic delay and dispersion to detect fMRI signal change without auditory interference. *Magnetic Resonance in Medicine.* 41, 13-20.

Ellis, A. W., Burani, C., Izura, C., Bromiley, A. and Venneri, A. (2006). Traces of vocabulary acquisition in the brain: Evidence from covert object naming. *NeuroImage.* 33:3, 958-968.

Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., Gutzwiller, F. and Lyrer, P. A. (2006). Epidemiology of aphasia attributable to first ischemic stroke – Incidence, severity, fluency, etiology and thrombolysis. *Stroke.* 37:6, 1379-1384.

Feigin, V. L., Lawes, C. M. M., Bennett, D. A. and Anderson, C. S. (2003). Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. *Lancet Neurology.* 2, 43-53.

Forster, K. I. and Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behaviour.* 12, 627-635.

Foygel, D. and Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language.* 43, 182-216.

Fridriksson, J., Morrow, K. L., Moser, D. and Baylis, G. C. (2006). Age-related variability in cortical activity during language processing. *Journal of Speech and Language Hearing Research.* 49, 690-697.

Friston, K. J., Josephs, O., Rees, G. and Turner, R. (1998). Nonlinear event-related responses in fMRI. *Magnetic Resonance in Medicine.* 39, 41-52.

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

Fu, C. H. Y., Morgan, K., Suckling, J., Williams, S. C. R., Andrew, C., Vythelingum, G. N. and McGuire, P. (2002). A functional magnetic resonance imaging study of overt letter verbal fluency using a clustered acquisition sequence: Greater anterior cingulated activation with increased task demand. *NeuroImage.* 17, 871-879.

Gehring, W. J. and Knight, R. T. (2000). Prefrontal-cingulate interactions in action monitoring. *Nature Neuroscience.* 3, 516-520.

Geschwind, N. (1970). The organisation of language and the brain. *Science.* 170, 940-944.

Gold, B. T. and Kertesz, A. (2000). Right hemisphere semantic processing of visual words in an aphasic patient: An fMRI study. *Brain and Language.* 73, 456-465.

Haarmann, H. J., Just, M. A. and Carpenter, P. A. (1997). Aphasic sentence comprehension as a resource deficit: A computational approach. *Brain and Language.* 59, 76-120.

Heiss, W. D., Karbe, H., WeberLuxenburger, G., Herholz, K., Kessler, J., Pietrzyk, U and Pawlik, G. (1997). Speech-induced cerebral metabolic activation reflects recovery from aphasia. *Journal of the Neurological Sciences.* 145:2, 213-217.

Heiss, W. D., Kessler, J., Thiel, A., Ghaemi, M. and Karbe, H. (1999). Differential capacity of left and right hemispheric areas for compensation of poststroke aphasia. *Annals of Neurology.* 45:4, 430-438.

Hillis, A. E. (2007). Aphasia – Progress in the last quarter of a century. *Neurology.* 69:2, 200-213.

Hogson, C. and Ellis, A. W. (1998). Last in, first to go: Age of acquisition and naming in the elderly. *Brain and Language.* 64, 146-163.

Howard, D., Patterson, K., Franklin, S., Morton, J. and Orchard-Lisle, V. Variability and consistency in picture naming by aphasic patients. *In:* FC Rose, ed. *Advances in neurology 42; progress in aphasiology*. New York: Raven Press, 1984.

Howard, D. and Patterson, K. (1992). Pyramids and Palm Trees: a test of semantic access from pictures and words. Thames Valley Test Company, Bury St Edmunds.

Howard, D. and Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology*. 20, 921-950.

Huang, J., Carr, T. H. and Cao, Y. (2001). Comparing cortical activations for silent and overt speech using event-related fMRI. *Human Brain Mapping*. 15, 39-53.

Huettel, S. A., Singerman, J. D. and McCarthy, G. (2001). The effects of aging upon the hemodynamic response measured by functional MRI. *NeuroImage*. 13, 161-175.

Huettel, S. A., Song, A. W. and McCarthy, G. (2004). *Functional Magnetic Resonance Imaging*. Sinauer Associates Inc., Massachusetts, USA.

Humphreys, G. W., Riddoch, M. J. and Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*. 51, 67-103.

Indefrey, P. and Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*. 92, 101-144.

Kan, I. P. and Thompson-Schill, S. L. (2004). Effect of name agreement on prefrontal activity during overt and covert picture naming. *Cognitive Affective Behavioural Neuroscience*. 4, 43-57.

Karbe, H., Thiel, A., Weber-Luxenburger, G., Herholz, K., Kessler, J. and Heiss, W. (1998). Brain plasticity in poststroke aphasia: what is the contribution of the right hemisphere? *Brain and Language*. 64, 215-230.

Kuahanen, M. L., Korpelainen, J. T., Hiltunen, P., Maatta, R., Mononen, H., Brusin, E., Sotaniemi, K. A. and Myllyla, V. V. (2000). Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke. *Cerebrovascular Diseases.* 10:6, 455-461.

Lambon-Ralph, M. A. (1998). Distributed versus localist representations: Evidence from a study of item consistency in a case of classical anomia. *Brain and Language.* 64, 339-360.

Laws, K. R. (2000). Category-specific naming errors in normal subjects: The influence of evolution and experience. *Brain and Language.* 75:1, 123-133.

Levelt, W. J. M., Praamstra, P., Meyer, A. S., Helenius, P. and Salmelin, R. (1998). An MEG study of picture naming. *Journal of Cognitive Neuroscience.* 10:5, 553-567.

Liu, T. T., Frank, L. R., Wong, E. C. and Buxton, R. B. (2001). Detection power, estimation efficiency, and predictability in event-related fMRI. *NeuroImage.* 13, 759-773.

MacDonald, A. W., Cohen, J. D., Stenger, V. A. and Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science.* 288, 1835-1838.

Martin, P. I., Naeser, M. A., Doron, K. W., Bogdan, A., Baker, E. H., Kurland, J., Renshaw, P. and Yurgelun-Todd, D. (2005). Overt naming in aphasia studied with a functional MRI hemodynamic delay design. *NeuroImage.* 28:1, 194-204.

McKenna, P. and Warrington, E. K. (1983). Graded naming test. NFER-Nelson, Windsor.

Medina, J. and Weintraub, S. (2007). Depression in primary progressive aphasia. *Journal of Geriatric Psychiatry and Neurology.* 20:3, 153-160.

Meinzer, M., Flaisch, T., Obleser, J., Assadollahi, R., Djundja, D., Barthel, G. and Rockstroh, B. (2006). Brain regions essential for improved lexical access in an aged aphasic patient: a case report. *BMC Neurology.* 6:28.

Moses M. S., Nickels, L. A. and Sheard, C. (2004). "I'm sitting here feeling aphasic!" A study of recurrent perseverative errors elicited in unimpaired speakers. *Brain and Language.* 89, 157-173.

Murata, Y., Sakatani, K., Katayama, Y. and Fukaya, C. (2002). Increase in focal concentration of deoxyhaemoglobin during neuronal activity in cerebral ischaemic patients. *Journal of Neurology, Neurosurgery and Psychiatry.* 73, 182-184.

Musso, M., Weiller, C., Kiebel, S., Muller, S. P., Bulau, P. and Rijntjes, M. (1999). Training-induced brain plasticity in aphasia. *Brain.* 122, 1781-1790.

Naeser, M. A., Martin, P. I., Nicholas, M., Baker, E. H., Seekins, H., Kobayashi, M., Theoret, H., Fregni, F., Maria-Tormos, J., Kurland, J., Doron, K. W. and Pascual-Leone, A. (2005). Improved picture naming in chronic aphasia after TMS to part of right Broca's area: An open-protocol study. *Brain and Language.* 93, 95-105.

Netz, J., Lammers, T. and Homberg, V. (1997). Reorganisation of motor output in the non-affected hemisphere after stroke. *Brain.* 120, 1579-1586.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia.* 9, 97-113.

Ojemann, G. A., Ojemann, J. G., Lettich, E. and Berger, M. (1989). Cortical language localisation in left-dominant hemisphere. *Journal of Neurosurgery.* 71, 316-326.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M. and Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *NeuroImage.* 14, 182-193.

Perani, D., Cappa, S. F., Tettamanti, M., Rosa, M., Scifo, P., Miozzo, A., Basso, A. and Fazio F. (2003). A fMRI study of word retrieval in aphasia. *Brain and Language.* 85:3, 357-368.

Persson, J., Sylvester, C-Y. C., Nelson, J. K., Welsh, K. M., Jonides, J. and Reuter-Lorenz, P. A. (2004). Selection requirements during verb generation: Differential recruitment in older and younger adults. *NeuroImage.* 23, 1382-1390.

Pineiro, R., Pendlebury, S., Johansen-Berg, H. and Matthews, P. M. (2002). Altered hemodynamic responses in patients after subcortical stroke measured by functional MRI. *Stroke.* 33, 103-109.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S. and Patterson, K. E. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review.* 103, 56-115.

Poeppel, D. and Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition.* 92, 1-12.

Price, C. J. and Friston, K. J. (1999). Scanning patients with tasks they can perform. *Human Brain Mapping.* 8, 102-108.

Rasmussen, T. and Milner, B. (1977). The role of early left-brain injury in determining lateralization of cerebral speech functions. *Annals of the New York Academy of Sciences.* 299, 355-369.

Rosen, H. J., Petersen, S. E., Linenweber, M. R., Snyder, A. Z., White, D. A., Chapman, L., Dromerick, A. W., Fiez, J. A. and Corbetta, M. (2000). Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology.* 55:12, 1883-1894.

Rossini, P. M., Altamura, C., Ferretti, A., Vernieri, F., Zappasodi, F., Caulo, M., Pizzella, V., Del Gratta, C., Romani, G.-L. and Tecchio, F. (2004). Does cerebrovascular disease affect the coupling between neuronal activity and local hemodynamics?. *Brain.* 127, 99-110.

Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology.* 8, 535-543.

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M. and Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain.* 129, 1371-1384.

Schwartz, M. F., Saffran, E. M., Bloch, D. E. and Dell, G. S. (1994). Disordered speech production in aphasic and normal speakers. *Brain and Language.* 47, 52-88.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S. and Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language.* 54:2, 228-264.

Seitz, R. J., Azari, N. P., Knorr, U., Binkofski, F., Herzog, H. and Freund, H. (1999). The role of diaschisis in stroke recovery. *Stroke.* 30, 1844-1850.

Seto, E., Sela, G., McIlroy, W. E., Black, S. E., Staines, W. R., Bronskill, M. J., McIntosh, A. R. and Graham, S. J. (2001). Quantifying head motion associated with motor tasks used in fMRI. *NeuroImage.* 14, 284-297.

Sharp, D. J, Scott, S. K., and Wise, R. J. S. (2004). Retrieving meaning after temporal lobe infarction: The role of the basal language area. *Annals of Neurology.* 56:6, 836-846.

Sharp, D. J., Scott, S. K., Mehta, M. A. and Wise, R. J. S. (2006). The neural correlates of declining performance with age: Evidence for age-related changes in cognitive control. *Cerebral Cortex.* 16, 1739-1749.

Siesjo, B. K. (1984). Cerebral circulation and metabolism. *Journal of Neurosurgery.* 60, 883-908.

Silkes, J. P., McNeil, M. R. and Drton, M. (2004). Simulation of aphasic naming performance in non-brain damaged adults. *Journal of Speech, Language and Hearing Research.* 47:3, 610-623.

Smith, S. M. (2001). Overview of fMRI analysis. In Jezzard, P., Matthews, P. M. and Smith, S. M. (Eds.), *Functional MRI: An Introduction to Methods* (pp. 215-227). Oxford University Press.

Snodgrass, J. G. and Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory.* 6:2, 174-215.

Szaflarski, J. P., Holland, S. K., Schmithorst, V. J. and Byars, A. W. (2006). fMRI study of language lateralization in children and adults. *Human Brain Mapping.* 27:3, 202-212.

Thomas, S. A. and Lincoln, N. B. (2006). Factors relating to depression after stroke. *British Journal of Clinical Psychology.* 45, 49-61.

Tsang, H. and Lee, T. M. C. (2003). The effect of ageing on confrontational naming ability. *Archives of Clinical Neuropsychology.* 18, 81-89.

Van Horn, J. D., Ellmore, T. M., Esposito, G. and Berman, K. F. (1998). Mapping voxel-based statistical power on parametric images. *NeuroImage.* 7, 97-107.

Vitkovitch, M. and Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. *Journal of Experimental Psychology: Learning, Memory and Cognition.* 17:4, 664-680.

Vitali, P., Abutalebi, J., Tettamanti, M., Danna, M., Ansaldo, A. I., Perani, D., Joanette, Y. and Cappa, S. F. (2007). Training-induced brain remapping in chronic aphasia: A pilot study. *Neurorehabilitation and Neural Repair.* 21:2, 152-160.

Warburton, E., Price, C. J., Swinburn, K. and Wise, R. J. S. (1999). Mechanisms of recovery from aphasia: evidence from positron emission tomography studies. *Journal of Neurology Neurosurgery and Psychiatry.* 66, 155-161.

Warrington, E. K. and James, M. (1999). Visual object and space perception battery. Thames Valley Test Company, Bury St Edmunds.