Inner Speech in Post-Stroke Aphasia: A Behavioural and Imaging Study

Sharon Geva

Wolfson College

Department of Clinical Neurosciences

This dissertation is submitted for the degree of Doctor of Philosophy

UNIVERSITY OF CAMBRIDGE

Declaration

The text in this dissertation does not contain more than 60,000 words, excluding figures, tables, appendices and bibliography. This dissertation was not submitted for a degree or diploma or other qualification at any other university. This dissertation is the result of my own work and includes nothing which is the outcome of work done by others or in collaboration except where specifically indicated in the text.

Summary

Patients with aphasia often complain that there is a poor correlation between the words they think (inner speech) and the words they say (overt speech). Previous studies show that there are some cases in which inner speech is preserved while overt speech is impaired, and vice versa. However, these studies have various methodological and theoretical drawbacks. In cognitive models of language processing, inner speech is described as either dependent on both speech production and speech comprehension, or on the speech production system alone. Lastly, imaging studies show that inner speech is correlated with activation in various language areas. However, these studies are sparse and many have methodological caveats. Moreover, studies looking at inner speech in stroke patients are rare.

This study examined inner speech in post-stroke aphasia using three different methodological approaches. Using cognitive behavioural methods, inner speech was characterised in healthy participants and stroke patients with aphasia. Using imaging, the brain structures which support inner speech were investigated. Two different methods were employed in this instance: Voxel based Lesion Symptom Mapping (VLSM) and Voxel Based Morphometry (VBM). Lastly, functional magnetic resonance imaging (fMRI) was used to study the dynamics of functional activations supporting inner speech production.

The study showed that inner speech can remain intact while there is a marked deficit in overt speech. Structural studies suggested an involvement of the dorsal language route in inner speech processing, together with systems supporting motor feedback and executive functions. Functional imaging showed that inner speech processing in stroke is correlated with compensatory peri-lesional and contra-lesional activations. Activations outside the language network might reflect increase in effort or attention, or the use of feed forward and feedback mechanisms to support inner speech production. These results have implications for diagnosis, prognosis and therapy of certain patients with post-stroke aphasia.

Acknowledgements

Many people helped me throughout these years, and I cannot thank them all enough. I would like to mention the ones who were most directly involved with my work.

Firstly, my supervisor, Liz Warburton, for her encouragement, enthusiasm and enormous professional and personal support, for giving me her love for knowledge, and for spending many hours discussing the mysteries of the human brain. I would also like to thank Jean-Claude Baron, for his help and advice throughout these years, and the members of the Stroke Research Group.

I would like to thank the patients and their families for their time and efforts.

My sincere thanks also go to Karalyn Patterson, who helped me enormously to understand the behavioural data, and Cathy Price and Jenny Crinion for their help with the imaging work.

I would also like to thank:

Simon Jones and Tulasi Marrapu, for many hours of analysis, statistics and programming and Diana Day for recruitment of patients.

The WBIC radiographers, especially Jon Campbell, Tracy Mattick, Sean Kearns, Rosna Hoque, Vicky Lupson and the rest of the WBIC staff.

Nicola Lambert, Helen Walker, Helen Palmer and Katherine Scantlebury for help with recruitment and testing of patients and for many hours of training they gave me, Sinead Stringwell for her help in the preparation of materials, and Sophie Bennet, for the help with the behavioural work.

The Pinsent-Darwin Fellowship, Wingate scholarship, The Cambridge Overseas Trust and B'nai Brith Scholarship for their generous funding.

And lastly, a big thanks to my parents, who always put our education on the top of their priorities. My parents started reading my home work the day I started learning the aleph-bet, and have not stopped since then. And my husband, Harith, who read my work, listened to my ideas and was supportive throughout the many up and down moments, and above all, always managed to give me this interesting outlook on my results, which only a medieval historian can.

Table of Contents

Appendices

List of Tables

List of Figures

Chapter 1: Introduction

Part I: Stroke, Aphasia and its Therapy

"No method of treatment is better than the principles on which it is based, and the search for principles should concern us no less than the immediate clinical situation" (Zangwill 1947).

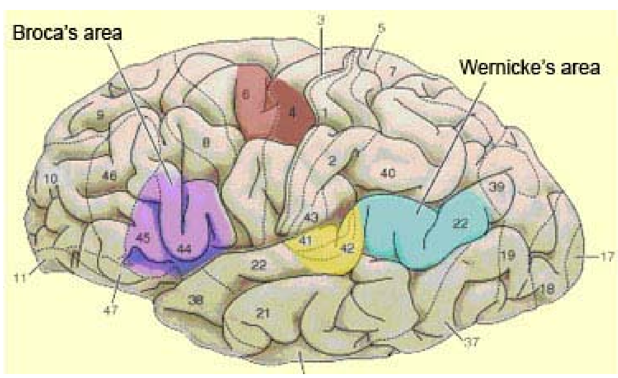
Stroke is the third most common cause of death and the single most common cause of disability in the UK (The Stroke Association). Estimates suggest that more than 20% of patients suffering a stroke develop aphasia, a loss or impairment of language function caused by brain damage (Benson and Ardila 1996). Longitudinal studies suggest that up to 50% of the survivors remain with a permanent deficit (Lazar and Antoniello 2008). Aphasia is characterised by communication problems, which include difficulties with talking, understanding spoken language, writing or reading (NHS Health Encyclopaedia on-line). Even everyday tasks such as making telephone calls, shopping, or having a conversation, can become impossible. There are currently approximately 250,000 people with aphasia in the UK (Heiss, Thiel et al. 2003, NHS Health Encyclopaedia on-line). There are still no reliable clinical predictors for recovery. Many factors have been shown to be linked to recovery, including type of aphasia, initial severity, and site and size of lesion. However, none can fully predict recovery across the wide spectrum of aphasia sympotmatology (Lazar and Antoniello 2008). Moreover, although most studies agree that aphasia therapy can contribute to recovery, the exact factors influencing it are unknown, and it is still unclear which patient will benefit from which therapy regime (Enderby and Emerson 1996; Robey 1998; Greener, Enderby et al. 2000; Bhogal, Teasell et al. 2003). It is therefore apparent that research into aphasia diagnosis and therapy is needed and can potentially benefit the large community of patients with post-stroke language deficits.

1. Etiology of Aphasia

1.1 Historical Background

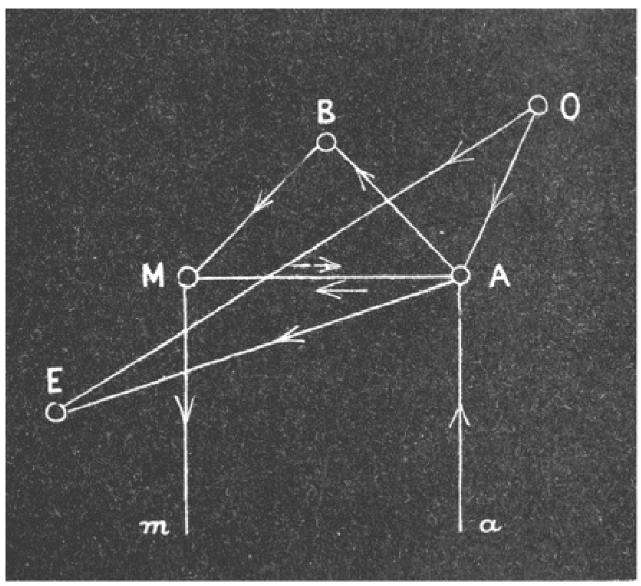
Almost every discussion of anatomy of aphasia starts with Paul Broca (1824-1880), the famous 19th century French physician, anatomist, and anthropologist. Broca is most famous for his work on speech production (1865). Broca examined a patient who had lost all speech production abilities, and when seen by him could only say the syllable 'Tan'. In a post-mortem autopsy of the patient's brain, Broca identified the inferior frontal gyrus of the left hemisphere as responsible for translating words to motor articulation commands. This area was soon to be called 'Broca's area' (figure 1.1; but see Dronkers, Plaisant et al. 2007, for MRI-based localisation of Broca's area). Tan's famous brain is preserved until today in the Musée Dupuytren in Paris. Broca is also accredited for defining the left hemisphere as the seat of language (Benson and Ardila 1996).

Figure 1.1. Broca's and Wernicke's areas in the left hemisphere (adapted from McGill University website: http://thebrain.mcgill.ca/). The numbers represent Brodmann Areas which will be discussed below.



Another early landmark in the history of aphasia is the publication by Carl Wernicke (1848-1905), a German physician, anatomist, psychiatrist and neuropathologist. Wernicke localised impairments of language comprehension to the left posterior superior temporal gyrus (1874), an area which later became known as 'Wernicke's area' (see figure 1.1). He defined the area as a sensory language area, responsible for mapping sounds to meaning. He also predicted that a disconnection between the anterior Broca's area and the more posterior Wernicke's area, would result in aphasia which is different from the ones described before, having production or comprehension deficits. Like Broca's area, Wernicke's area was recently redefined, both functionally and structurally (Wise, Scott et al. 2001). Later, Lichtheim (1845-1928) developed a model of language processing (1885). This model described and predicted various language deficits, therefore laying the foundations for later cognitive models of language processing. It also tried to relate the described language centres to anatomical structures (Compston 2006). Figure 1.2 presents a copy of Lechtheim's model together with descriptions of the anatomical localisation.

Figure 1.2. Lichtheim's model of speech processing (1885) includes a word-representations centre located in the left temporal lobe (A); a motor images centre located at the inferior frontal convolutions of both hemispheres and descending towards the corticobulbar pathways (M); a modulating centre distributed across the entire hemisphere (B); a processor which transmit acoustic information (a); a motor articulation pathway (m); a visual centre for reading (o); and a centre to execute writing (E). The connection between 'A' and 'M' passes through the insula, while the left anterior frontal lobe connects 'O' to 'A'.



Dejerine (1849-1917) examined patients with reading and writing disturbances. He determined that alexia with agraphia, impairment in both reading and writing, is caused by a lesion to the angular gyrus, where visual images of letters and words are stored. He also suggested that pure alexia, an isolated reading impairment, is a disconnection syndrome, caused by a lesion to the visual cortex, or specifically to areas connecting visual stimuli with reading centres (Basso 2003).

At the beginning of the 20th century criticisms of the 'localisation approach' grew in Europe and scholars and practitioners such as Sigmund Freud (1856-1939), John Hughlings Jackson (1835-1911) and Pierre Marie (1853-1940), argued for an holistic approach for understanding language, strongly opposing the idea that modular centres in the brain process different aspects of language (Benson and Ardila 1996). Constantin von Monakow (1853-1930) introduced the term ‘diaschisis’ to describe abnormal function in areas distal from the focus of the lesion. He attempted to resolve the debate between the localisation and holistic approaches by arguing that both diaschisis and the localised lesion contribute to the clinical picture seen in patients (Stein, Brailowsky et al. 1995). Lashley’s studies of animal models (1929) showed that the amount of cortex damaged, but not the location of the damage, influences the resulting learning behaviour and level of impairment. Although the idea was developed in relation to memory and learning, it had a large influence over the scientific community and in aphasia, supported the holistic view (Benson and Ardila 1996). In Russia, Alexander Luria (1902-1977) took a more balanced stance by arguing that language acts are complex cognitive functions which depend on many distributed areas of the cortex. He suggested that language areas act as association cortices, binding information sourced from various parts of the cortex (Luria 1964).

Many consider Norman Geschwind (1926-1984) to be the recent revivalist of the localisation approach (Benson and Ardila 1996; Basso 2003; Poeppel and Hickok 2004). Geschwind restated the importance of language areas but also described various association syndromes, caused by damage to the white matter tracts connecting these areas (Basso 2003). His students, followers and other scholars in the area, put an enormous effort into understanding the functional anatomy of language, an area which is today one of the most prolific topics in brain sciences.

In more recent years, there have been some major developments which influenced both theoretical and methodological aspects of aphasia research. First, Broca's and Wernicke's areas are not seen as homogenous areas, any longer (for example, see the review on Broca's area by Keller, Crow et al. 2009). Secondly, many other brain areas have been highlighted as contributors to language processing (Poeppel and Hickok 2004). Regarding experimental paradigms and methodology, there have been two major advances. First, the evolution of neuroimaging techniques has enabled researchers to investigate both the structure and the function of the living human brain. Secondly, a more sophisticated understanding of language per se, gave rise to the development of psychological, linguistic and neuroanatomical models of language, which are used to guide research (Dronkers 2000; Wise 2003; Poeppel and Hickok 2004).

It is impossible to attempt to summarise all the current models and contemporary debates in the area of language processing and language anatomy (Poeppel and Hickok 2004). Therefore, a comprehensive review of the neural correlates of language is outside the scope of this introduction. The next section will give a general review of the current understanding of the functional anatomy of language. This will enable a better understanding of the anatomical studies presented throughout this thesis.

In general terms, there are two main ways to define brain structures anatomically. One is by referring to Brodmann Areas (BA) and the other by defining the relative location of the structure in each of the four cortical lobes (frontal, parietal, temporal and occipital). Moreover, some areas have names which are related to their discoverer or suggested function (for example, Wernicke's area or the Visual Word Form Area). In the following paragraphs the different references will be, as much as possible, combined. Appendix 1 presents the anatomy of the cortex showing the various labels used in the literature today.

1.2 Functional Anatomy of Language

1.2.1 Functional Anatomy of Speech Comprehension

Comprehension of spoken language begins with the processing of auditory signals in bilateral primary auditory cortex (A1), located in the supratemporal plane (BA41 and 42, also known as *Heschl's gyri*). This stage of processing is not specific to language. In the next stage, acoustic–phonetic analysis (phonetic discrimination) takes place in the middle-posterior region of the superior temporal gyrus (STG) – lateral to A1, an area which is considered to be an auditory association cortex (reviewed in Price 2000; Boatman 2004; Demonet, Thierry et al. 2005). This process involves both hemispheres although there might be a slight bias towards the left (Hickok and Poeppel 2007). Subsequently, phonological processing takes place, whereby auditory input is mapped onto existing phonological representations. It involves the anterior middle STG, ventral and dorsal portions of the posterior STG, and there is suggested involvement of the inferior frontal lobe (reviewed in Martin 2003; Boatman 2004; Hickok and Poeppel 2007; Saur, Schelter et al. 2009).

Many scholars suggest that at this point the processing stream diverges into two pathways, paralleling the two pathways for visual processing (Hickok and

Poeppel 2007; Saur, Kreher et al. 2008). The first pathway is the ventral (rostral) route which passes through the uncinate fasciculus, towards ventral and rostral prefrontal cortex (BA44/45). This route is thought to support the understanding of word meaning (reviewed in Wise 2003). The next stage in this route is therefore mapping of sound to meaning (phonological-semantic mapping). This process is probably relatively distributed (Martin 2003) and takes place in the anterior lateral superior temporal lobe, where processing of complex semantics in comprehension of sentences (Crinion, Warburton et al. 2006; Warren, Crinion et al. 2009) and perception of intelligible speech (reviewed in Wise 2003) take place, as well as in the posterior temporal and inferior parietal areas (BA22/39) (reviewed in Boatman 2004). Others (Price 2000; Friederici 2002; Martin 2003; Hickok and Poeppel 2007; Saur, Schelter et al. 2009) suggest that phonological-semantic mapping also occurs in the posterior middle temporal gyrus (MTG) and inferior temporal sulcus. Moreover, some argue that this process happens bilaterally, an idea which is supported by findings from semantic dementia (Lambon-Ralph, Pobric et al. 2009), cognitive impairments following Herpes Simplex virus (Noppeney, Patterson et al. 2007) and other studies (reviewed in Patterson, Nestor et al. 2007).

The second processing route, the caudal (dorsal) route goes through the arcuate fasciculus towards the dorsolateral prefrontal cortex (DLPFC, BA46), and is involved in analyzing word sound (reviewed in Wise 2003; Saur, Kreher et al. 2008). It integrates auditory and motor information, to allow language acquisition (feed forward mechanisms), monitoring of overt speech (feedback mechanisms, Hickok and Poeppel 2007), and repetition (Saur, Kreher et al. 2008). The dorsal stream therefore sends information through the inferior parietal area, to the articulatory network, which includes the inferior frontal gyrus (IFG/Broca's area/BA44-45) (Hickok and Poeppel 2007; Saur, Kreher et al. 2008), the anterior insula (Hickok and Poeppel 2007) and the premotor area in the middle frontal gyrus (MFG/BA6) (Wilson, Saygin et al. 2004; Hickok and Poeppel 2007; Saur, Kreher et al. 2008). Beyond the word level, sentence comprehension requires syntactic processing, which is thought to occur in the anterior left STG (Martin 2003; Friederici, Bahlmann et al. 2006), BA44, and the left frontal operculum (Friederici 2002; Friederici 2006) and also involves the dorsal route (Friederici, Bahlmann et al. 2006).

1.2.2 Functional Anatomy of Speech Production

The first step in speech production is semantic access, in which the speaker decides what concepts he or she wants to convey in speech. This process takes place during various tasks such as picture naming, spontaneous speech and verb generation. Some models suggest that the following process is *lemma* selection[1], in which the speaker chooses the appropriate lexical entry to express the semantic idea. This lexical entry specifies the syntactic information related to the wanted word and acts as a link between the semantic system and the *lexeme*, the phonological form of the word (Levelt 1993). However, the distinction between the lemma and the lexeme has been debated in the literature, with Caramazza and Miozzo (1997; 1998), for example, strongly arguing against it. In their studies, they suggested that data which traditionally was interpreted as supporting the lemma/lexeme distinction can be interpreted otherwise. They bring neuropsychological and other evidence to support a more parsimonious model, in which lemmas are not needed. A recent meta-analysis of imaging studies of speech production (Indefrey and Levelt 2004) does not distinguish the 'lemma selection' from other processes, and therefore fails to contribute to this debate. According to this meta-analysis, the two processes, semantic access and lemma selection, both involve the middle section of the MTG (Indefrey and Levelt 2004). Schwartz et al. (2009) have shown that the left anterior temporal lobe is involved in translating semantics (concepts) into words, during the process of word retrieval. These findings are in agreement with some studies of speech comprehension showing that semantic processing takes place in the middle MTG, and with some classical language models (for example, Dell 1986; Levelt 1989) which assert that at this processing level, comprehension and production share a common mechanism.

The next process, phonological retrieval[2], is probably widely distributed with several areas involved, including the right supplementary motor area (SMA), left anterior insula and the left posterior superior and middle temporal gyrus (Indefrey and Levelt 2004), left anterior operculum, left posterior basal temporal lobe (BA37) and the midline cerebellum (Price and Friston 1997). Lastly, preparation and execution of articulation employs mainly the bilateral motor system (Wise, Greene et al. 1999;

---

[1] Also known as lexical retrieval. Lemma contains the lexical aspects of a given word (Levelt et al. 1999).

[2] According to Levelt's model (1999) this process includes three sub-processes: morphological encoding, phonological encoding and phonetic encoding, the latter is also referred to as preparation of the articulatory code.

Indefrey and Levelt 2004) but also other regions, including the right posterior IFG, left orbital gyrus, bilateral posterior lingual gyri and the right posterior medial temporal fusiform (Indefrey and Levelt 2004), Broca's area (Hillis, Work et al. 2004) and the insula (Wise, Greene et al. 1999).

1.2.3 Functional Anatomy of Reading

Reading begins with visual processing which takes place in the bilateral ventral extrastriate cortex (Schlaggar and McCandliss 2007). According to the classical Dual-Route model of Paap and colleagues (Paap, McDonald et al. 1987; Paap and Noel 1991), the next stage in reading involves translating graphemes[3] to phonemes, which allows access to the so called 'phonological input lexicon', in one of two possible routes. The first route is the lexical route, dedicated to reading frequent regular as well as irregular words by means of whole word recognition. The second, sublexical, route supports the reading of new words and non-words, by utilising direct grapheme to phoneme translation. This model is the subject of ongoing debate among researchers. It was later revisited and modified, mainly by Coltheart and colleagues who developed the Dual-Route Cascaded (DRC) Model (Coltheart, Curtis et al. 1993). Conversely, Seidenberg and McClelland (1989) challenged the Dual-Route model and proposed a more distributed, connectionist-like, model of reading. Other researchers suggest that a direct orthography to semantic route exist as well, not requiring access to phonology (for example, see review by Houghton and Zorzi 2003).

Some suggest that these lexical and sublexical routes can be distinguished anatomically and mapped onto ventral and dorsal systems, respectively (reviewed in Price 2000; Pugh, Mencl et al. 2000; Schlaggar and McCandliss 2007). The ventral, lexical, system involves the left occipito-temporal region, and is responsible for word identification based on stored, known, word forms (Thuy, Matsuo et al. 2004). This area has been subsequently named the Visual Word Form Area (VWFA, Cohen, Dehaene et al. 2000). However, some argue against naming the left occipitotemporal area the VWFA, because of practical and theoretical reasons, the most important of them probably being that the region does not demonstrate such functional specificity, as the name might imply (Price and Devlin 2004). Some studies show that this route also involves the left medial lingual gyrus, which is not specific to reading, but rather

[3] Graphemes are the minimal units in the writing system of a specific language. In languages which use alphabet, such as English, graphemes often correspond to letters.

involved in global visual processing (Price 2000). The dorsal, sublexical, network involves the supramarginal gyrus, angular gyrus, and superior posterior temporal cortex. This temporo-parietal system is responsible for rule-based analysis of orthography to phonology (reviewed in Pugh, Mencl et al. 2000; Schlaggar and McCandliss 2007). Others have shown that lexical and sublexical reading impairments are not correlated with any one specific lesion, therefore emphasizing a more distributed network for spelling (Rapcsak, Beeson et al. 2009). However, in the same study it was also found that anterior perisylvian lesions (in the regions of the posterior inferior frontal gyrus/Broca's area and the pre-central gyrus) were more associated with phonological dyslexia (Rapcsak, Beeson et al. 2009).

Based on their study of literate and illiterate adults, Carreiras et al. (2009) have suggested that the dorsal route is more important in the early stages of learning to read, when reading is heavily based on direct translation of orthography to phonology, and in adults, in languages where orthography is relatively transparent, such as Spanish. The ventral route is more prominent in proficient readers, and in languages with deeper orthographies (such as English).

Indefrey and Levelt (2004) distinguish core processes of language from lead-in processes, which are task specific. For example in picture naming, object recognition would be an example of a 'lead-in process', being task specific, and not part of the language production system. In reading, the processes discussed above would be the lead-in processes. Hence in reading, after the written orthographic form allowed access to the phonological form, the processing stream 'joins' the language comprehension system described above, accessing semantics and therefore allowing reading comprehension, or joining the language production system, allowing reading aloud.

1.2.4 Functional Anatomy of Writing

The first stages of writing engage the language production system, described above. Whether writing to dictation, or writing spontaneously, one has to access the phonological form of the word. While in speech production this phonological form is then translated into an articulatory code, in writing it needs to be translated into letters (Nolan and Caramazza 1982). However, other suggest that orthography can be accessed directly from semantics, without the need of phonological mediation (Piras and Marangolo 2004; Miceli and Capasso 2006). Some models suggest that as with reading, writing can be accomplished through two different routes: a lexical route and a sub-lexical route, mirroring the dual-route model for reading (Houghton and Zorzi 2003). However, these routes are probably not completely independent and can influence each other during writing (Rapp and Goldrick 2006; White, Abrams et al. 2008). Accessing the orthographical output lexicon through the lexical route takes place in the left angular gyrus (for a review of studies of dementia, stroke and healthy participants, see Houghton and Zorzi 2003), extending into the inferior parietal gyrus (Booth, Cho et al. 2007), while the sublexical route involves left premotor area, extending into Broca's area (BA6/44) (Omura, Tsukamoto et al. 2004). Henry et al. (2007) have demonstrated that patients with perisylvian lesions tend to have symptoms of phonological dysgraphia, while patients with extra-sylvian lesions, especially around the superior and inferior parietal lobe, tend to have symptoms associated with lexical dysgrpahia. These authors argue against a strict dual-route model and a stringent lesion-symptom mapping, since they could not find any one lesion site which is associated with lexical or sublexical spelling impairment. Instead, they emphasise the importance of both phonological and semantic feedback in the process of writing (Henry, Beeson et al. 2007). As with the reading models, the main criticisms to the dual-route model of spelling come from connectionist models in which sound to orthography translation occurs in multilayered single route, with back propagation (Brown and Loosemore 1994; Bullinaria 1994; Olson and Caramazza 1994).

The next stage in writing involves accessing the letter images, which are thought to be stored in the anterior superior parietal lobe (Katanoda, Yoshikawa et al. 2001). Lastly, motor images related to writing specific letters are stored in premotor areas, located in the posterior parts of the middle and superior frontal gyri, the former being known as Exner’s area (Katanoda, Yoshikawa et al. 2001). Exner’s area is named after Siegmund Exner (1846-1926), an Austrian physiologist who described the relation between lesions to this region and writing impairments in 1881 (Benson and Ardila 1996).

2. Diagnosis of Aphasia

2.1 Diagnostic Tests

In the 19th century reports on cases of aphasia were mainly descriptive, based on patient observation (Benson and Ardila 1996). The first well-known aphasia battery was published by Head, in 1926. Although Head (1926) tested healthy people with his battery, he did not publish his results, and therefore the battery cannot be seen as officially normalised. Nine years later, Weisenburg and McBride (1935) published their test for aphasia. This test was an influential landmark because it introduced the concept of normalisation into aphasia testing and was comprehensive (Byng, Kay et al. 1990). In subsequent years many aphasia tests were developed, but their popularity only rose in the 1970s and 1980s (Howard, Swinburn et al. 2010). Currently used tests vary in their language, content, and theoretical approach. By far the most commonly used tests are those based on the modality approach, which characterises the patient's symptoms in the four language modalities: comprehension, production, reading and writing (Bruce and Edmundson 2010). Commonly used tests of this kind include the Minnesota Test for Differential Diagnosis of Aphasia (MTDDA, Schuell 1965), the Boston Diagnostic Aphasia Examination (BDAE, Goodglass and Kaplan 1972) the Western Aphasia Battery (WAB, Kertesz 1982), the Porch Index of Communicative Ability (PICA, Porch 1971), widely used in the USA, the Aachen Aphasia Test (AAT, Huber, Poeck et al. 1984), widely used in German speaking countries, and its variations in other languages including Dutch, Italian, English and Thai. Lastly, for British English speakers the Comprehensive Aphasia Tests (CAT, Swinburn, Porter et al. 2004) is a well-designed, widely available, test.

These tests vary considerably in their content, aims and underlying theoretical approach. For example, the WAB and BDAE are built upon 'localisation approaches' and therefore seek to relate the aphasic symptoms to brain lesions. The MTDDA, on the other hand, relies on the idea that all language disorders are a result of a single impairment to 'learned auditory patterns'. The WAB and BDAE also seek to classify individual patients according to specific aphasia syndromes, as do the AAT and the MTDDA. The CAT gives up the idea of 'syndromes' altogether, and instead it seeks to define the underlying impairments of each patient with regard to different components of the language system and various linguistic variables known to influence it, such as word frequency and imageability (Howard, Swinburn et al. 2010). The PICA is aimed at providing a measurement of recovery thus helping in giving prognosis. Other tests are aimed at diagnosing only specific subtypes of aphasia. For example, the Boston Naming Test (BNT, Kaplan, Goodglass et al. 2001) and the Graded Naming Test (Warrington and McKenna 1983) both examine symptoms of anomia. The Token Test (Derenzi and Vignolo 1962) examines auditory comprehension. Lastly, there are tests which aim at defining the functional abilities and disabilities of the patient. These include the Functional Communication Profile (Sarno 1969) and the Communicative Abilities in Daily Living (CADL, Holland 1980), among others.

A rise in interest in neuropsychological 'information processing' models (Byng, Kay et al. 1990), brought about the development of tests based on these models, such as the Psycholinguistic Assessment of Language Processing (PALPA, Kay, Coltheart et al. 1992) and the CAT. These developments are based on the view that the use of such models allow a better understanding of the patient's symptoms and a better planning of a suitable rehabilitation programme (Bruce and Edmundson 2010). Neuropsychological models were developed based on language studies of normal participants. They describe the processes employed by the normal brain when performing specific language tasks, and include the processors and the connections linking them, portrayed as boxes and arrows, respectively (for examples of such models see Ellis and Young 1988; Levelt 1999; Martin 2003). During the diagnosis process, patients' deficits are localised to a specific component of the system. Some neuropsychologists hold the 'assumption of modularity', which claims that the system's components can be individually damaged while the rest of the system is kept intact. However, some components, so they surmise, might be intact but still not get their normal input, therefore they are unable to perform their role in the process (Byng, Kay et al. 1990). This is in fact a behavioural description of diaschisis, mentioned above.

2.2 Main Problems with Current Diagnostic Tests

There are various problems with current tests of aphasia. Firstly, tests which relate language dysfunction to brain lesion are becoming increasingly irrelevant since advanced neuroimaging techniques can localise lesions much more reliably (Bruce and Edmundson 2010). Some tests are not based on established cognitive models of language, thus risk disregarding relevant linguistic factors which influence aphasia (Byng, Kay et al. 1990; Bruce and Edmundson 2010). Byng (1990) suggests that classifying patients into syndromes gives little information about the underlying impairment, making such 'classifying tests' inadequate for rehabilitation planning. This problem was already pointed out more than one hundred years ago, by Curnow, who stated that 'the tendency to appear exact by ignoring the complexity of the factors is the old failing of our medical history' (cited in Hughlings Jackson 1879, p. 309). The issue of classification attracts wide interest from clinicians and researchers alike. Recently, Ardila suggested a new classification method for the aphasias (Ardila 2010), a suggestion which brought about a lively debate, with some authors (for example, Marshall 2010) arguing that classification does not contribute to neither therapy nor research. Authors also question the sensitivity (Howard and Hatfield 1987; Byng, Kay et al. 1990), reliability (Biddle, Watson et al. 2002) and predictive validity (Bruce and Edmundson 2010) of some aphasia batteries. Another criticism is that some important aspects of language function are not adequately assessed by aphasia batteries (Simmons-Mackie, Threats et al. 2005). This latter issue underlies some of the motivation for the current research, and together with the other criticisms brought up here, will be addressed later.

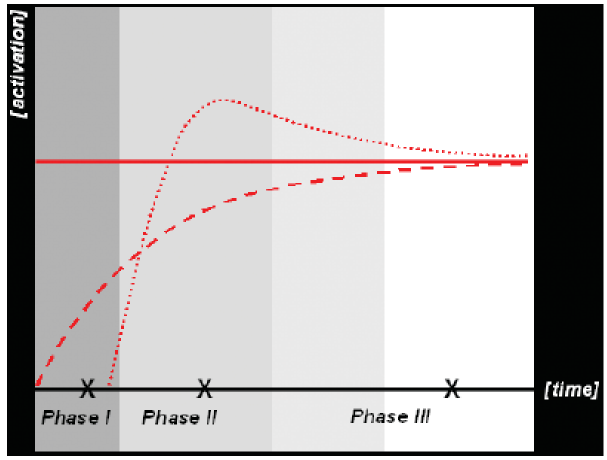
3. Recovery from Aphasia

3.1 Spontaneous Recovery after Stroke

It is well established that stroke patients experience some spontaneous recovery of behaviour during the first few months after the stroke. While the behavioural manifestations of this recovery are well documented, there is an ongoing debate over the neural correlates of this recovery, specifically, what levels of reorganisation accompany recovery and whether it involves ipsi- or contra-lesional activity in the brain (Rijntjes 2006; Cramer 2008). Moreover, scholars disagree about the time scale in which spontaneous reorganisation can occur. In a recent review, Cramer (2008) argued that in general, spontaneous reorganisation lasts for 3 months. However, he also noted that while cognitive deficits might show spontaneous improvement beyond this time limit, motor deficits are less likely to spontaneously improve after 3 months. Moreover, some studies show spontaneous changes in behaviour months and even years after the stroke (Cramer 2008).

Plastic reorganisation in the adult brain after stroke can occur in response to central injury at different levels. This has been investigated widely in the motor system, starting with early positron emission tomography (PET) studies (Frackowiak, Weiller et al. 1991; Weiller, Chollet et al. 1992; Weiller, Ramsay et al. 1993). Studies have demonstrated reorganisation within the same area (peri-lesional reorganisation, Nudo, Wise et al. 1996), at remote cortical areas (in humans and in squirrel monkeys, Cramer, Nelles et al. 1997; Frost, Barbay et al. 2003), and even in the unaffected, contra-lateral hemisphere (Cramer, Nelles et al. 1997). Regarding language, there is an ongoing discussion whether recovery-related reorganisation happens only in peri-lesional areas (Furlan, Marchal et al. 1996; Iglesias, Marchal et al. 1996; Price, Mummery et al. 1999; Warburton, Price et al. 1999) or in contra-lesional areas as well (Weiller, Isensee et al. 1995; Cappa, Perani et al. 1997; Calvert, Brammer et al. 2000; Winhuisen, Thiel et al. 2005). Some suggest that contra-lesional activation is a maladaptive reaction (Naeser, Martin et al. 2004; Price and Crinion 2005) and can even be a predictor of poor recovery (Kurland, Naeser et al. 2004). This activation might be a result of loss of inhibition to the contra-lateral hemisphere, also known as transcallosal disinhibition (Price and Crinion 2005). Some studies demonstrated that recovery is characterised by an initial stage in which recovery-related activation can be seen in contra-lesional areas (Fernandez, Cardebat et al. 2004; Xu, Zhang et al. 2004; Saur, Lange et al. 2006), followed by recovery-related activation in peri-lesional areas (Fernandez, Cardebat et al. 2004; Saur, Lange et al. 2006). While some of those studies looked at patients in the chronic or acute stage, a recent study (Saur, Lange et al. 2006) tried to document the recovery process by testing patients at three time points: acute, subacute and chronic. Using a language comprehension task, the authors showed that brain activation, as seen in fMRI, was restricted to a small area in the left IFG in the acute stage. At the subacute stage, activation was vast and bilateral, while at the chronic stage, activation was 'normalised', and did not differ from controls. This included activation in Wernicke's and Broca's areas, the left temporal gyrus, the SMA and the right IFG. Behavioural improvement was observed both at the subacute and chronic stages. The authors concluded that early bilateral activation is followed by left lateralised activation in rehabilitating patients with aphasia. Their model is shown in figure 1.3.

Figure 1.3. Model of three phases of language recovery after stroke. Diagrammed activation of controls (___), left language areas (- - -) and right language areas of patients with aphasia (. . .). Crosses (X) indicate time of examinations. Adapted from Saur et al. (2006).



Price et al. (2001) introduced the term *dynamic diaschisis* to refer to the process which results in abnormal activation seen in unaffected brain regions after stroke. This abnormal activity is explained to be a result of abnormal input from affected areas, and is thought to be task specific. If the task normally involves the affected brain areas, then unaffected areas might not receive the normal input and would therefore show abnormal activation. However, these areas can still receive normal input from other brain regions, hence, in other contexts (other tasks), the activation can be normal (Price, Warburton et al. 2001). This might explain the discrepancies between different studies, with some showing only peri-lesional activation during recovery, while others showing distant and even contra-lesional activation in recovering stroke patients.

3.2 Aphasia Rehabilitation

3.2.1 Aphasia Therapy

Alongside spontaneous recovery, most patients in the UK and many in other countries receive some kind of therapy. Therapy approaches differ in the assumptions that they make regarding the patients' impairments and abilities. In general, approaches can be divided into the restoration and compensation approaches. The *restoration approach* concentrates on restoring language abilities (Benson and Ardila 1996). It is based on the idea that for many patients with aphasia, some linguistic knowledge is preserved but access to it is impaired (Benson and Ardila 1996). Therefore, therapy is aimed at restoring access to this knowledge (Whitworth, Webster et al. 2005). A strong support for the idea that representations are spared, comes from observations of inconsistency in impairment in specific items and responsiveness to cues. *Inconsistency in impairment* means that on different occasions patients show difficulties with different words rather than always having problems with the same words. In this case one can argue that if the representation of a word is impaired, then the patient will have consistent difficulty with this word. However, if the deficit is in general word access mechanisms, then patients would show an impairment for different words at different times (Warrington and Shallice 1979; Basso 2003), which is indeed the case with many patients (Howard, Patterson et al. 1984). *Responsiveness to cues* means that patients are more likely to perform a language task successfully after receiving a cue. This further supports the notion that the representation of the word exists but access to it is impaired. If the representation itself was impaired, the patient would not be able to benefit from cues. Lastly, many patients show spared automatic speech in the absence of voluntary speech (Lum and Ellis 1994; Lum and Ellis 1999). These patients are able to count, recite, sing and perform other automatic language tasks, while showing a profound impairment in voluntary spontaneous speech (Benson and Ardila 1996). This as well has been taken as evidence for spared linguistic representations (Basso 2003).

Contrary to the restoration approach, the *compensation* approach assumes that the linguistic knowledge is lost and therefore the only way to treat the patient is by developing a compensatory means of communication such as gestures, drawing, sign language and the like (Howard and Hatfield 1987; Whitworth, Webster et al. 2005). Many studies have looked at the effectiveness of different therapy techniques. For example, studies on the effectiveness of phonological treatment (Miceli, Amitrano et al. 1996; Harding and Pound 1999; White-Thomson 1999; Spencer, Doyle et al. 2000; Franklin, Buerk et al. 2002; Nickels 2002), semantic treatment (Hillis 1989; Marshall, Pound et al. 1990; Pring, Hamilton et al. 1993; Hillis and Caramazza 1994), and both types of treatments (Nettleton and Lesser 1991; Doesborgh, van de Sandt-Koenderman et al. 2004) demonstrated the contribution that such treatments have to rehabilitation. However, Ledoreze et al. (1994) found semantic therapy to be ineffective. In a meta-analysis of studies of aphasia treatment, it was found that; first, treating patients is beneficial, secondly, the earlier the treatment, the greater the benefits are for the patient, and lastly, treatment should be delivered for at least two hours weekly, in order to achieve positive outcomes (Robey 1998). However, the author stresses that no study complied with all criteria for controlled clinical trials. The author adds that: '…having settled the general issue (that treatment is beneficial, SG) it is important now to expend resources in testing focused hypothesis… (e.g. certain patients improve more with the administration of a particular treatment protocol than with the administration of a different treatment protocol on a single schedule under homogenous circumstances)' (ibid. p. 183). Conversely, another review of aphasia therapy after stroke did not find that speech therapy has beneficial effects beyond spontaneous recovery (Greener, Enderby et al. 2000). The authors state that the major finding of their review is that there are only very few randomised trials, and those studies which do adhere to the selection criteria of good studies have only a small sample size (Greener, Enderby et al. 2000). A review by Nickels (2002) looked at studies examining individual therapy (case studies) for word finding difficulties. Nickels joins Robey's conclusion that studies of speech and language therapy often have methodological and analytical flaws. She also concludes that speech and language therapy can be beneficial for some individuals. However, she argues that we are still unable to predict which therapy will benefit which patient and to what extent. Lastly, the author argues that there is wide agreement that patients with different impairments should be treated differently (Nickels 2002). Finally, Wisenburn and Mahoney (2009) conducted a meta-analysis of therapy for word finding difficulty. They have found that overall treatment is efficient and evident after three months, but they could not determine whether one type of therapy (phonological, semantic or mixed) is more beneficial than the others. Interestingly, they have also found that gains can still be made in the chronic stage (Wisenburn and Mahoney 2009). The differences between the reviews might be partially attributable to the fact that studies evaluating intervention for aphasia use control groups of different kinds and often the 'standard therapy' (which the control group receives) is poorly or vaguely described, making it impossible to evaluate or replicate the results. An attempt to define 'standard therapy' has recently been made in the area of motor rehabilitation (Hunter, Crome et al. 2006). However, similar work has not been done in the area of speech and language therapy.

Future efforts to improve and understand speech therapy should focus on finely grained diagnoses of the patients' impairments, aiming at matching each patient with the most beneficial therapy technique for his or her own deficits and abilities. These will probably require large scale, multi-centre trials.

3.2.2 The Neural Correlates of Aphasia Therapy

As explained above, speech and language therapists, as well as neuropsychologists, commonly diagnose patients with aphasia by using detailed language assessment based on cognitive language models. Neurologists, on the other hand, use short bedside tests, as well as information about brain damage and other abnormalities, such as reduced or delayed blood flow, to construct their diagnosis and prognosis (Warlow, Dennis et al. 1996). Combining these two approaches can potentially lead to a more accurate diagnosis and prognosis of patients with aphasia. Studies looking at the neural correlates of language rehabilitation after stroke are aimed at doing exactly this. These have the potential to inform and guide therapy for language by revealing the pattern of activation of related neural mechanisms (Cornelissen, Laine et al. 2003). However, the methodological difficulties and pitfalls in conducting such research are numerous, and as a result, meaningful data sets are sparse. For example, in a recent study patients were treated with intense group therapy, aimed at improving both language production and comprehension. In an ERP session, patients' cortical responses during a lexical decision task were measured. The authors report an increase in cortical signal after treatment (Pulvermuller, Hauk et al. 2005). However, the lack of a 'control' condition, in the shape of a control group, control task, control items or a control session, makes it difficult to draw definite conclusions. Belin et al. (1996) used fMRI to explore changes in cortical activation after therapy for language production. Therapy targeted speech production but during the imaging session patients performed a repetition, rather than a production, task. Furthermore, again there was no adequate 'control' group, no imaging session before therapy, and large heterogeneity within the group of patients. Taken together, the ability to draw conclusions from this study remains limited. Another study looked at two chronic aphasia patients and their brain activation patterns before and after therapy (Cherney and Small 2006), but like previous studies, it had no control group, multiple baselines, or any other method which will reliably show that improvement in general language performance was therapy-dependent. Lastly, using a different methodological approach, Peck et al. (2004) studied 3 patients, who received three 2-week long treatments (5 days/week). Two patients received an *intention treatment*, aimed at improving initiation in general, and speech initiation is particular. Another patient received an *attention treatment*, aimed at increasing attention abilities, which in turn might improve speech production. During scanning, patients were asked to produce a single word in a presented category. Time-to-Peak (TTP) of the hemodynamic response was calculated in 4 right hemispheric regions which showed reliable activation in all patients. Pearson correlation coefficient between speech performance and change in TTP was very high (0.82, p=0.05). However, selecting brain regions and calculating correlations based on the same dependent data often results in high correlations (Vul, Harris et al. 2009). This is an example of a specific source of statistical bias which was recently termed 'double dipping' (Kriegeskorte, Simmons et al. 2009). To summarise, it is clear that designing and conducting studies into rehabilitation related brain activations is difficult and challenging. However, there are some well-designed studies which examined the neural correlates which support language rehabilitation. These are shortly reviewed in the next section (for a recent review see Meinzer and Breitenstein 2008). All studies presented are also summarised in table 1.1.

3.2.2.1 Case Studies of Therapies for Single Word Production

Most studies looking at the neural correlates of aphasia therapy assessed patients' response to therapies targeting single word production, and have focused on case-studies rather than groups of patients. The first study to look at brain activation pre- and post- intensive intervention is the one by Leger et al. (2002). RC, a patient with chronic and severe speech production deficits was treated with intensive (1h/day) therapy for two weeks. The therapy targeted articulation problems. After therapy, RC showed improved naming for both treated and untreated items. Therapy-related activations were found in peri-lesional areas, including Broca's area and the superior posterior part of the left supramarginal gyrus (BA40), both known to be related to phonological processing (Leger, Demonet et al. 2002). In a Magnetoencephalography (MEG) study, Cornelissen et al. (2003) studied three chronic patients with moderate anomia. The patients underwent semantic treatment for word production. Language assessment and brain imaging were completed twice before treatment and twice after, to ensure that observed changes are treatment-related. During scans patients performed a delayed picture naming task. For two patients, therapy resulted in a lasting, but non-generalised, improvement of naming and signal intensification in the left inferior parietal cortex, known to be related to phonological processing. This signal was specific to trained items. For the third patient, behavioural benefits from therapy generalised to semantically related items but were small and did not have a long-lasting effect. This patient had a reduction of signal in the left inferior parietal cortex. Lastly, right hemispheric signals did not change after therapy (Cornelissen, Laine et al. 2003). In another study, two chronic patients with moderate non-fluent aphasia and anomia were treated using the intention and attention treatments mentioned above. These therapies are based on the findings that intention (to move/speak) and attention precede execution, and take place in the right hemisphere. Increasing attention/intention can therefore improve performance. The first patient was expected to benefit from the intention therapy only. And indeed, the patient showed the expected results: his performance improved with therapy, and brain activation shifted to become more lateralised to the right hemisphere. However, that was not the case for the second patient. This patient was expected to benefit only from one type of therapy. However, she showed improved performance after both types of therapy. Moreover, her brain activation pre-treatment was already right lateralised, and therefore an inversion of the lateralisation was not expected. The authors interpret the results in light of the different lesions the two patients had (Crosson, Moore et al. 2005). Fridriksson et al. (2006) examined three patients with word finding difficulty before and after treatment. Therapy comprised of intensive exposure to a set of words chosen by the patient. One patient showed no improvement in performance, but in the other two patients, improvement in performance was correlated with increased cortical activity in the peri-lesional area. For one patient, improvement in performance was correlated with increased cortical activity in the left temporal pole (a peri-lesional area) and right inferior parietal lobe, as seen with fMRI. Yet for another patient, improved performance was correlated with left parietal lobe (a peri-lesional area) activation. The latter also showed increased activation in the frontal poles, bilaterally, and in the anterior cingulate gyrus, which might be related to inhibition and attention processes, respectively (Fridriksson, Morrow-Odom et al. 2006). The same group treated three further patients; two with Broca's aphasia and one with conduction-type aphasia. Therapy included semantic and phonological treatments in an intensive cross-over design. The task during the fMRI scan was overt naming. Activation post treatment was seen in non-language areas, such as those related to attention and working memory. The authors argue that these activations reflect compensatory mechanisms which support language rehabilitation in the context of lack of ‘spared’ language areas (Fridriksson, Moser et al. 2007). Meinzer et al. (2007) looked at cross-language generalisation[4] after therapy, in a 35-year-old bilingual stroke patient. The patient received Constraint Induced Therapy (CIT) for German, and was scanned before and after therapy. Behavioural analysis showed a clear improvement in naming of trained items in German, but without any generalisation into his second language, French, or to untrained items. The authors observed increase activation in both peri-lesional and contra-lesional areas. This change in activation was not observed for the non-trained items and language (Meinzer, Obleser et al. 2007). Two recent studies (Meinzer, Flaisch et al. 2006; Vitali, Abutalebi et al. 2007) are the only ones to directly correlate correct responses and errors with brain activation before and after therapy. In the first study, a chronic patient with word finding difficulty and Wernicke’s aphasia was treated with CIT for 3h/day, over ten consecutive days. Therapy resulted in improved naming. An increased activation during picture naming in the right IFG was found when correct trials were compared to trials in which the patient produced paraphasias or neologisms. Similar activation, together with right hemispheric subcortical activation, was also found when comparing erroneous responses pre-therapy, with correct responses post-therapy, on the same items (Meinzer, Flaisch et al. 2006). The second study examined two patients with severe chronic aphasia (Vitali, Abutalebi et al. 2007). Both patients improved on naming after a phonological therapy. The first patient was a young man suffering from a closed head injury. Before therapy, neural activity during production of errors was seen in the left IFG. After therapy, training-related activation was found in the spared left Broca’s area and in the left supramarginal gyrus. The second patient, on the other hand, had a large ischemic stroke encompassing most of the left hemispheric language system. Before therapy, activation during the production of errors was found in the right hemispheric homologue of Broca’s area, as well as in other right hemispheric areas. After therapy, training-related activation was seen mostly in right hemispheric areas, including the Broca’s homologue (Vitali, Abutalebi et al. 2007). Alas, the authors did not directly contrast the activation seen before and after therapy, or that seen for correct and incorrect responses.

[4] Generalisation is the process in which items that are not being directly treated, benefit from therapy as well (Nickels and Best, 1996).

3.2.2.2 Therapy for Syntactic Processing

Only one study so far attempted to characterise brain activation underlying therapy for syntactic processing (Wierenga, Maher et al. 2006). Two patients with word retrieval difficulty and impairments in sentence formation were given *syntactic mapping treatment* over the stretch of two months. This therapy technique is aimed at restoring the ability to produce meaningful and grammatically correct sentences. The imaging task consisted of silent generation of sentences. Both patients benefited from the treatment, but only the second patient showed a generalisation effect. While the first patient showed significant increase in Broca's area activation after treatment, the second patient showed a decrease in activation in the inferior frontal sulcus. The authors interpret the increase in activation as a support for normal performance. The decrease in activation seen for the second patient is interpreted as increased efficiency following practice (Wierenga, Maher et al. 2006). However, it should be noted that the analysis has the same methodological problem as in Peck et al. (2004). Namely, ROIs were chosen based on activation and then analysed further. This creates a selection bias, in which dependent data is analysed as independent, violating the assumptions of the model used for analysis (Kriegeskorte, Simmons et al. 2009). Moreover, interpretation of results is limited by the fact that no behavioural measurement was recorded during imaging.

3.2.2.3 Group Studies

The above studies conducted case-studies. In order to reach meaningful generalisable clinical conclusions, group studies are clearly essential. Using MEG, Meinzer et al. (2004) looked for correlation between changes in performance and abnormal slow waves, indicative of dysfunctional information processing. A group of 28 patients with aphasia benefited from the intensive therapy used, but a correlation between this improvement and brain activation was not found. While about half of the patients showed a decrease in slow waves, the other half showed an increase in these brain waves. The authors do not explain these puzzling results. However, it was also found that the magnitude of change was correlated with the magnitude of improvement in language ability (Meinzer, Elbert et al. 2004). 11 of the 28 patients who participated in the study were chosen for a follow-up study. Regions of interest (ROIs) for analysis of fMRI data were defined as those areas of abnormal slow waves seen previously in MEG during rest. All patients received 10 days of intensive CIT.

Improved picture naming of trained items was related to, and only to, the decrease in slow waves in the defined ROI, but not in any other control regions (Meinzer, Flaisch et al. 2008). A main strength of this study is that the ROIs were defined for each patient individually, using independent data, therefore reducing the risk of overlooking significant results due to variability between patients. Contrary to most studies, this therapy produced a generalisation effect; patients' performance improved for non-trained items. However, this improvement in performance was not correlated with the measured brain activation. This brings up the question of whether brain mechanisms supporting naming of untrained items are different from the ones supporting naming of trained items. The authors speculate about the behavioural processes which led to the observed generalisation, but the question regarding brain mechanisms underlying generalisation remains open. In a PET study, 10 patients with language production deficits underwent treatment for lexical retrieval, from which they benefited as a group. A group of healthy controls (who unfortunately were not age-matched to the patients), underwent a process of learning words in a foreign language, imitating the intervention process the patients went through. Both groups were scanned while performing a picture naming task, before and after intervention. Both groups showed treatment related activation in the right insular and right inferior frontal cortices, which the authors explain as being related to pronunciation and word retrieval. Compared with controls, patients showed therapy related activations in the orbital frontal cortex (BA11) and the medial frontal cortex (BA 10). These are interpreted as motivational and emotional related activation (Raboyeau, De Boissezon et al. 2008), although alternative cognitive explanations might be appropriate as well (Meinzer and Breitenstein 2008). Unsurprisingly, the opposite comparison (controls>patients) yielded no significant activations (the lesioned areas were excluded from analysis, because otherwise the control group would have inevitably showed higher activation than patients in the areas where patients have lesions). Lastly, Richter et al. (2008) demonstrated that intensive CIT over a 10 days period, benefited 16 chronic aphasia patients with speech production impairments. The patients were scanned before and after treatment, while engaged in reading and stem completion tasks. Before therapy, patients exhibited greater activation than controls in right hemispheric areas, explained as representing less effective strategies for language processing. Interestingly, activation in these right hemispheric regions before treatment was predictive of therapy outcome. This means that after therapy, activation in these areas increased for patients with small behavioural improvements and decreased for patients with good outcome (Richter, Miltner et al. 2008). The importance of the latter study is that it used functional imaging as a predictive tool for recovery. Previously, similar methodologies have been applied to post-stroke motor recovery (for example, see Cramer, Parrish et al. 2007; Marshall, Zarahn et al. 2009), but this study by Richter et al. is the first to do so in the area of language rehabilitation after stroke. It is important to note however, that these studies of motor recovery differ from the aphasia study in a couple of aspects. Firstly, the motor studies correlated activation at the acute stage with outcome at the chronic stage, while in the aphasia study, fMRI was conducted at the chronic stage only. The only study which paralleled the motor studies in the area of aphasia is one by Saur et al. (2010) in which brain activation in the acute stage (12 days post-stroke, on average) was correlated with general language outcome at the chronic stage (6 months post-stroke). Secondly, the motor studies looked at general outcome, while the aphasia study examined therapy related outcome. A study which combines acute imaging with specific therapy does not, to the best of my knowledge, exist in the literature. Table 1.1 summarises the studies presented in this section.

Table 1. Studies of neural correlates of language therapy after stroke. Studies appear in the same order as in the text.

Case Studies of Therapies for Single Word Production

Study	**Patients**	**Therapy**	**Imaging time points**	**Imaging technique**	**Task during imaging**	**Outcome Measurement**
Leger et al. (2002)	1 chronic patient	Intensive therapy for articulation	Once before and once after therapy	fMRI	Picture naming, and picture and written word rhyme judgement	Picture naming
Cornelissen et al. (2003)	3 chronic patients with moderate anomia	Semantic treatment for word production	Twice before and twice after therapy	MEG	Picture naming	Picture naming
Crosson et al. (2005)	2 chronic patients with moderate non-fluent aphasia and anomia	Intention and attention therapy	Once before and once after therapy	fMRI	Category–member-generation task	Picture naming
Fridriksson et al. (2006)	3 chronic patients suffering from word finding difficulty	Intensive exposure to a set of words chosen by the patient	Once before and once after therapy	fMRI	Picture naming	Picture naming
Fridriksson et al. (2007)	Chronic patients: 2 with Broca's aphasia and 1 with conduction aphasia	Intensive semantic and phonological therapy	Twice before therapy, twice after first therapy session and twice again after second therapy session. This is a cross-over design.	fMRI	Picture naming	Picture naming
Meinzer et al. (2007)	1 chronic bilingual patient	Constrain-induced-therapy (CIT) for German	Once before and once after therapy	fMRI	Picture naming	Naming of trained and untrained pictures in German and French
Meinzer et al. (2006)	1 chronic patient with word finding difficulty and Wernicke's aphasia	CIT	Once before and once after therapy	fMRI	Picture naming	Errors and correct responses during picture naming
Vitali et al. (2007)	2 chronic patients with severe chronic aphasia	Phonological therapy	Once before and once after therapy	fMRI	Picture naming	Errors and correct responses during picture naming

Therapy for Syntactic Processing

Study	Patients	Therapy	Imaging time points	Imaging technique	Task during imaging	Outcome Measurement
Wierenga et al. (2006)	2 chronic patients with word finding difficulty and impairments in sentence formation	Syntactic mapping treatment	Once before and once after therapy	fMRI	Silent generation of sentences	Sentence production

Group Studies

Study	Patients	Therapy	Imaging time points	Imaging technique	Task during imaging	Outcome Measurement
Meinzer et al. (2004)	28 chronic patients	Therapy for spoken word production, based on the patients' deficit	Once before and once after therapy	MEG	Rest	Behavioural profile according to the Aachen-Aphasia-Test (AAT)
Meinzer and Flaisch (2008)	11 of the 28 patients who participated in Meinzer et al. (2004).	CIT	Once before and once after therapy	fMRI	Picture naming	Picture naming
Raboyeau et al. (2008)	10 chronic patients with language production deficits	Treatment for lexical retrieval	Once before and once after therapy	PET	Picture naming	Picture naming
Richter et al. (2008)	16 chronic aphasic patients with speech production impairments	CIT	Once before and once after therapy	fMRI	Reading and stem completion tasks	Global score for behavioural improvement based on spontaneous speech, the Token Test, and auditory and semantic comprehensibi lity of speech.

3.2.2.4 Summary

Recent years have seen a growth in well designed studies, looking into the neural correlates of language rehabilitation after stroke. These studies shed some light on the basic neural mechanisms which support recovery of language production in the chronic stages of stroke. Taken together, these studies demonstrate that intensive therapy can improve language functions even in the chronic stages, that effects of generalisation are rare, and that behavioural improvement is linked to changes in neural activity. The published studies differ widely in the locus of therapy-related activations they find: some reveal activation which is confined to peri-lesional areas, others demonstrate vast contra-lesional activations. The exact roles of peri- versus contra-lesional activations, then, remain to be determined. Many questions still remain open. For example; are there differences in the neural mechanisms supporting recovery for treated versus untreated words/sentences? Can we, based on neural correlates and brain damage, fit patients with the best type of rehabilitation technique? How can generalisation be achieved and what are the neural substrates supporting it? These and other questions still remain to be answered using well-designed studies with multiple participants.

4. Theoretical Background of the Current Research

Taken together, it is evident that further development of rehabilitation techniques for language deficits after stroke need to be based on an enhanced understanding of patients' deficits and more accurate behavioural and imaging descriptions. These will then allow improved matching of each patient with the most suitable therapy technique.

This research explored a novel approach to diagnosis and rehabilitation of speech production impairments after stroke. The study paralleled promising developments that have already been made in research studying the rehabilitation of patients suffering from post-stroke motor deficits. While these patients cannot produce a movement, they are sometimes able to imagine such a movement (Sharma, Pomeroy et al. 2006). Such 'motor imagery' is suggested to promote motor performance, through the enhancement of brain activation (Jeannerod 1995). Some studies suggest that the use of motor imagery can be beneficial for stroke patients with motor deficits (for a recent review see Sharma, Pomeroy et al. 2006). The use of mental imagery has also been suggested for the treatment of visual neglect (Luaute, Halligan et al. 2006). One study showed that two patients who were treated with a mental imagery therapy for left visual neglect, as a result of right hemispheric damage, showed treatment-related benefits immediately after the treatment as well as six months post treatment (Smania, Bazoli et al. 1997).

Drawing from this approach, this research explored the characteristics of speech imagery (i.e. ‘inner speech’) in post-stroke aphasia. Inner speech is the ability to ‘talk to oneself… develop an auditory-articulatory image of speech without uttering sound’. It also refers to ‘…the objectively measurable ability to appreciate the auditory-articulatory structure of speech irrespective of its meaning’ (Levine, Calvanio et al. 1982, p.391). In simple words, inner speech is the ability to talk to yourself in your head and listen to what you are saying - a cognitive action which we all engage in every day. The next part will describe previous studies of inner speech and its depiction in current cognitive models.

Part II: Inner Speech: Past Research and Current Models

1. Historical Perspectives

Inner speech[5], the ability to speak silently in your head, has been suggested to play an important role in memory (Baddeley and Hitch 1974), reading (Corcoran 1966), thinking (Sokolov 1972), language acquisition (Vygotsky 1962), language comprehension (Blonskii 1964), auditory hallucination in schizophrenia (Brown 2009), and even in consciousness and self-reflective activities (Morin and Michaud 2007).

In ancient and medieval times inner speech was the domain of philosophers, who were trying to understand the 'language of the mind' and the connection between thought and language. The first to investigate inner speech using the methodology of experimental psychology were Egger (1881) and Ballet (1886). By using introspection they tried to understand the relation between inner speech and thought and by doing so they also brought about an outburst of experimental work on inner speech (reviewed in Sokolov 1972). The next wave of experimental work started in 1899 by Curtis, and continued into the beginning of the 20th century. Researchers were now interested in understanding the connection between inner speech and articulation (De Bleser and Marshall 2005). Curtis (1899) recorded laryngeal movement during silent recitation and silent reading of verses. His results showed that 75% of participants made motor movements during the production of inner speech. The studies that followed showed that while some participants, in some conditions, produce motor movements during silent speech, others do not (reviewed in Sokolov 1972). More recent experimental work shows that participants vocalise more as the difficulty of reading material increases, and have lower comprehension scores when prevented from vocalising or even sub-vocalising[6] (Hardyck and Petrinov 1970).

One cannot discuss inner speech without referring to Lev Vygotsky, the famous Russian developmental psychologist, who explored the concept of inner speech, and its relation to the development of language. Vygotsky (1962) argued that

[5] Inner speech is also referred to as internal speech, covert speech, speech imagery, silent speech, auditory imagery, self-talk, subvocal speech, internal dialogue/monologue, subvocalization and self-verbalisation.
[6] The authors define subvocal speech as inner speech which is accompanied by laryngeal movements.

young children have no inner speech and therefore they can only think out loud. With the acquisition of language, speech becomes increasingly internalised. Mature inner speech, he argued, is different from overt speech in that it lacks the complete syntactic structures available in overt speech, and its semantics is personal and contextual rather than objective. Vygotsky's main critic was Blonskii, who argued that infants do have inner speech, and that it is vital for speech acquisition. Blonskii (1964) suggested that when infants hear other people speak, they repeat the sounds in their head, thus producing inner speech, which later allows them to produce overt speech. Moreover, by examining results from experimental work he suggested that repetition using inner speech is a prerequisite for overt repetition even in adults. However, Blonskii agreed with Vygotsky that adult inner speech is fundamentally different from overt speech in that it is syntactically impoverished and semantically different. Recent work by Oppenheim and Dell (2008) suggest that inner speech is phonetically impoverished in comparison to overt speech, because inner speech lacks some of the phonetic components present in overt speech, or, because the comprehension system fails to detect the full range of phonetic features of the produced inner speech. In both cases, the result is that the speaker has access to less phonetic information with inner speech than with overt speech. This might impair performance if only inner speech is available. However, Corcoran (1966) showed that participants automatically access phonetics in inner speech. In his seminal study he described the use of inner speech in reading processes. In this study, participants were asked to silently read a paragraph and cross out the letter ‘e’ wherever it occurred. Readers were less likely to detect a silent ‘e’ than a pronounced ‘e’. This phenomenon was explained by the fact that readers normally access phonology during silent reading using inner speech. When looking for the letter ‘e’, readers listen to their inner speech and notice the letter whenever they hear the sound ‘e’. Therefore, when the letter ‘e’ is silent, they are less likely to detect it (Corcoran 1966).

With the recent rise of neuroimaging, researchers turn to PET and fMRI to enhance the understanding of questions previously explored by philosophers and cognitive psychologists. Many brain imaging studies of language use covert response (inner speech) as the preferable response mode, apparently assuming that overt and inner speech differ only in the articulatory motor component present in overt speech. However, such an assumption has not been established (Huang, Carr et al. 2002; Gracco, Tremblay et al. 2005; Shuster and Lemieux 2005), and in fact, some studies provide reasons to doubt this assumption. Direct comparisons between conditions of overt and inner speech indicate that although they yield overlapping brain activation, the two conditions also produce activations in other regions of the brain, reflecting distinct *non-motor* cognitive processes (Ryding, Bradvik et al. 1996; Barch, Sabb et al. 1999; Palmer, Rosen et al. 2001; Huang, Carr et al. 2002; Indefrey and Levelt 2004; Shuster and Lemieux 2005; Basho, Palmer et al. 2007).

2. Current Models of Inner Speech

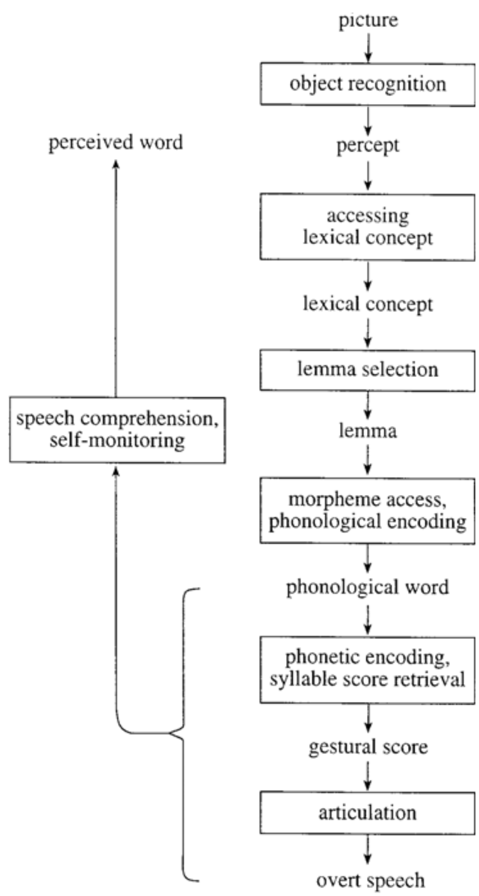
Within the framework of current cognitive models, the one which gives the most attention to inner speech is the working memory model (Baddeley and Hitch 1974) which includes the phonological loop. The phonological loop is divided into two components: the phonological store and the articulatory control process (Salame and Baddeley 1982; Vallar and Baddeley 1984). The phonological store holds verbal information for a short period of time in which the information will decay unless actively rehearsed (Conrad 1964; Baddeley 1966). The articulatory control process is an active component that allows rehearsal of words - properties of which were initially defined by observing the word length effect (Baddeley, Thomson et al. 1975). These two components are influenced by different types of input and interference (Buchsbaum and D'Esposito 2008). Most cognitive models of language production and language comprehension put relatively less emphasis on inner speech. However, it is important to consider inner speech in the context of these models as well, since it is essentially a linguistic process which might be active irrespective of working memory. Moreover, while the working memory model was initially developed within the memory discipline, and was seen as independent from the language system, some researchers have since argued that language and verbal working memory cannot be fully dissociated, because they are supported by common mechanisms (Martin, Saffran et al. 1996; Hulme, Roodenrys et al. 1997; Martin and Saffran 1997; Martin, Lesch et al. 1999; Page, Madge et al. 2007; Acheson and MacDonald 2009).

In Levelt et al.'s (1999) model of speech the systems for language production and comprehension are described as partially separate[7] (figure 1.4) - a view that is supported by others (for example see Martin, Lesch et al. 1999; Martin 2003).

[7] Separation occurs at the phonological level. The lexical and semantic levels are common to production and comprehension.

However, it is surmised that information can flow between the systems. Levelt et al. suggest that, for monitoring purposes, the phonological representation for production (*phonological word*) is perceived by the comprehension system to create inner speech.

Figure 1.4. A model of language production and comprehension (adapted from Levelt, Praamstra et al. 1998). The phonological word can be transferred into the comprehension system to create inner speech, for monitoring purposes, as described above. Overt speech can be perceived by the comprehension system as well, supporting another source of feedback, as will be further described below.



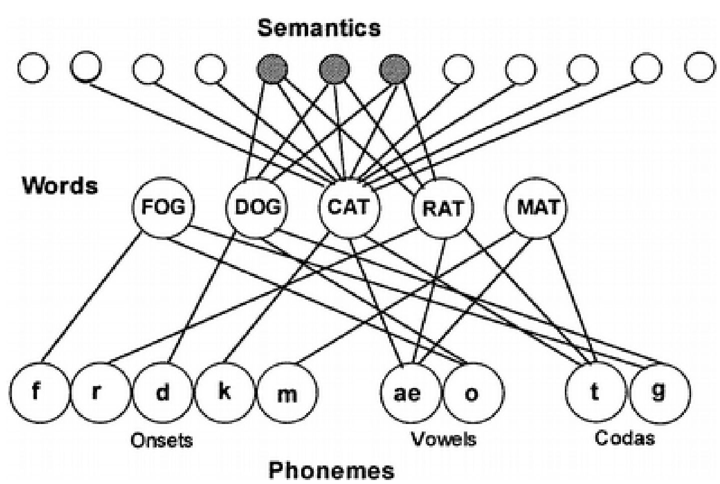
Monsell (1987), in his account for inner speech, describes the same process as a flow of information from the output phonological store into the input phonological store[8]. According to these models, inner speech is dependent on both the production and the comprehension systems, and inner and overt speech can to some degree be independent from one another. Thus, one might be damaged while the other remains intact. To evaluate this theory, Ozdemir et al. (2007) have examined the influence of the 'uniqueness point' of a word on monitoring for the presence of specific phonemes in internal speech. A word's uniqueness point is the place in its sequence of phonemes at which it deviates from every other word in the language; hence, it makes the word

[8] Information can flow in the other direction as well: from the input phonological store to the output phonological store, as can be seen in cases of phonological priming by non-words during naming (Levelt et al. 1999) and non-word repetition (Martin et al. 1999).

'unique'. The uniqueness point is known to influence speech perception, but not speech production. Ozdemir et al. (2007) reported that 'uniqueness point' influenced inner speech, a finding which they take to support the idea that inner speech is dependent on the speech comprehension system.

In contrast, Vigliocco and Hartsuiker (2002) argued that in the presence of overt speech, inner speech is processed by the production system alone. A study by Huettig and Hartsuiker (2009) showed differences in eye movements driven by internal speech compared to eye movements driven by one's own overt speech and overt speech produced by others. This finding seems to support the idea that inner speech is not processed by the comprehension system, but Vigliocco and Hartsuiker also suggest that in the absence of overt speech, what they describe as 'conscious inner speech' can access the comprehension system. Therefore, according to Vigliocco and Hartsuiker, inner speech can be processed either jointly by the production and comprehension systems, or by the production system alone, and the actual mechanism employed depends on the presence or absence of overt speech. Howard and Franklin (1990) agree that inner speech can be processed in these two ways, and suggest that the 'choice' between them is task-dependent. Lastly, in the connectionist model of Dell (1986), feedback connections within the production system can support the monitoring system without involving the comprehension system (figure 1.5).

Figure 1.5. Dell's interactive two-step model of naming (Dell, Schwartz et al. 1997). All connections in the model are bidirectional, allowing the production of inner speech within this one system. This representation of one integrated system for both speech comprehension and production with bidirectional connections bare similarities to that proposed by Vigliocco and Hartsuiker (2002).

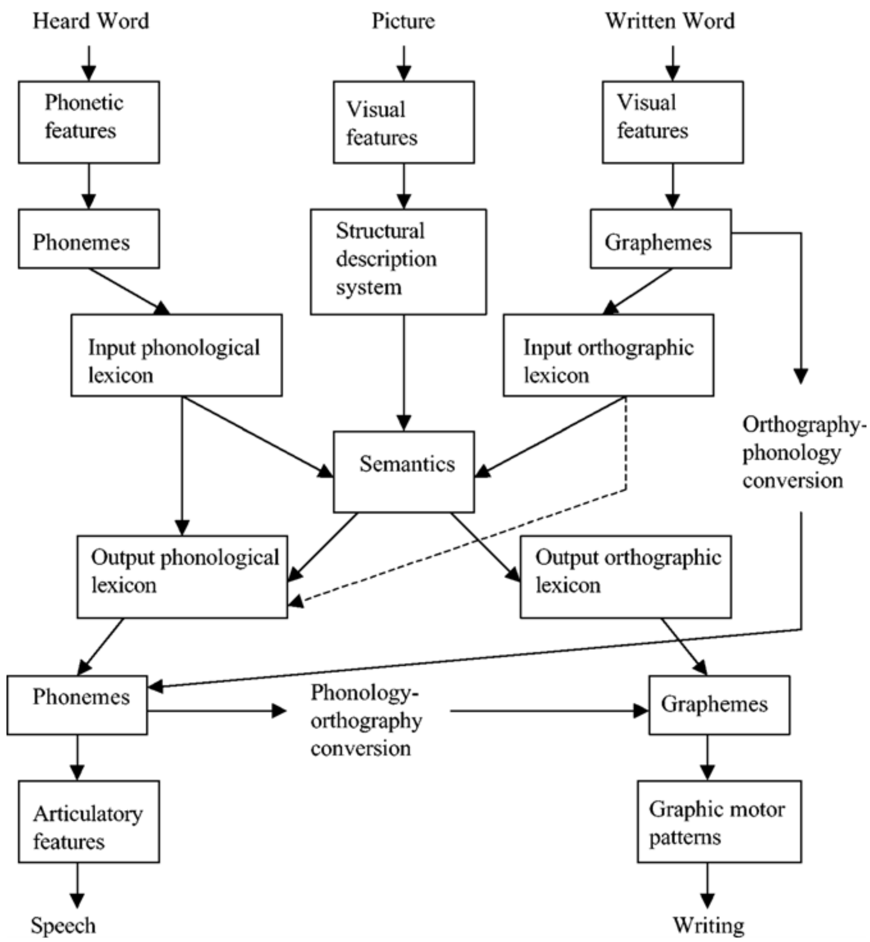


This idea is supported by data from Marshall et al. (1998) and others (for example Nickels and Howard 1995) who showed that in some cases of aphasia, comprehension can be relatively intact while monitoring of one's own errors is severely impaired and vice versa.

Since inner speech is important in reading processes, its depiction in current models of reading should be considered as well. However, current models of reading do not describe inner speech specifically, so one has to combine models of reading, with those of speech production, to create a full picture of inner speech production. In their classical model, Paap and colleagues (Paap, McDonald et al. 1987; Paap and Noel 1991) describe a model with two routes for reading. This model was briefly described above, and is explained again here. The first route, the lexical route, is dedicated to reading frequent regular, as well as irregular words. The process involves whole word recognition, followed by accessing the phonological lexical entry, which leads on to semantic access. The second, the sublexical route, supports the reading of new words and non-words, by utilising direct grapheme to phoneme translation. Both routes converge at the 'phoneme units' node, which in many ways parallels the input phonological store described above. Combining the models of speech production and reading aloud, one can argue that the phoneme unit can either transfer information to overt speech production, or be used to create inner speech. This model of reading, then, does not lend support to one model of inner speech over the other. As mentioned before, the main challenges to the dual route model were presented by Seidenberg and McClelland (Seidenberg and McClelland 1989) with their distributed, connectionist-like, model of reading. Importantly though, both models have something in common: they both claim that the propagation of information is bidirectional; both 'top down' and 'bottom up'. Top down processing can therefore support the reading of words by activation from the 'phonological lexicon' and 'semantic lexicon' to the 'orthographic lexicon', according to Coltheart et al. (Coltheart, Curtis et al. 1993), or from 'meaning' and 'phonology' nodes to 'orthography' if using Seidenberg et al.'s terminology (Seidenberg and McClelland 1989).

Martin (2003) reviews the different models and studies of oral speech production, auditory speech comprehension and reading. The review is based on Ellis and Young's (1988) early model. Although their model has subsequently been modified, their basic claim, that word production processes converge, is widely accepted (Martin 2003), and can be clearly visualised in the model below (figure 1.6). This model gives a framework for integrating some of the various models presented in this introduction. Specifically, it describes the convergence of models which emphasise the separation between speech production and speech comprehension.

Figure 1.6. A comprehensive model of language processing, based on Ellis and Young's model (1988). Adapted from Martin (2003).



3. The Current Research

More than sixty years ago, Oliver Zangwill wrote: 'No method of treatment is better than the principles on which it is based, and the search for principles should concern us no less than the immediate clinical situation' (Zangwill 1947, p.7). The goal of this present research was, therefore, to 'search for principles' underlying a potential technique for diagnosing and treating aphasia using inner speech.

This research project investigated the behavioural manifestations of inner speech and its underlying neural mechanisms. From the review of historical and contemporary studies of inner speech, it is clear that inner speech production has various functions, processing stages and access mechanisms. Different studies focus on different processing stages: from the initial stage in which thought is translated into words, to the final stage of articulation which may or may not be related to inner speech. Accessing inner speech can be either spontaneous, as is the case when we think to ourselves in words, or stimuli driven, giving rise to inner speech produced, for example, during reading. Spontaneous inner speech also supports error monitoring, and it is still debated whether this type of inner speech is based on the production system alone or on both the production and the comprehension systems.

In this study inner speech was defined as the ability to create an internal representation of the auditory word form, and to apply computations or manipulations to this representation. Participants were asked to create inner speech based on a given written stimuli. It is acknowledged that this type of inner speech might differ in some respects from the one used for error monitoring or spontaneous thinking. The following questions were investigated:

1. Do stroke patients with aphasia have impairments of inner speech, and if they do, what is the relation between inner and overt speech abilities? Inner speech was measured by asking subjects to determine whether two written letter strings rhyme or sound the same. Overt speech scores represent participants' ability to read single words aloud. This question was addressed by using neuropsychological tests and is presented in chapters 2 and 3.

2. Which brain regions are necessary for the production of inner speech and do these areas overlap with those areas supporting overt speech production or speech comprehension? This question was answered by examining structural changes in the brain, using Voxel-based Lesion Symptom Mapping (VLSM), and Voxel Based Morphometry (VBM). These studies are presented in chapters 4 and 5.

3. Which areas support inner speech production in the healthy brain and in post-stroke aphasia? What are the differences in brain activation between patients with different lesions and different behavioural profiles? What are the differences in brain activation between healthy controls and stroke patients? These questions were explored by using functional Magnetic Resonance Imaging (fMRI), and the results are presented in chapter 6.

These three research questions are explored in the next chapters. Each chapter includes an introduction specific to the research question, a description of the methods and materials used, a detailed description of the results and a short discussion. The last chapter presents a comprehensive discussion of all of the data presented in this thesis, as well as a discussion of the main research caveats and possible directions for future research. Appendix 2 describes the exclusion and inclusion of patients in different parts of the study.

Chapter 2
Inner Speech in Normal Language Processing

"The soul when thinking appears to me to be just talking – asking questions of itself and answering them, affirming and denying. And when it has arrived at a decision, either gradually or on a sudden impulse, and has at last agreed, and does not doubt, this is called its opinion. I say, then, that to form an opinion is to speak, and opinion is a word spoken - I mean, to oneself and in silence, not aloud or not to another."

Plato, Theaetetus

1. Introduction

The first stage in this research was to characterise the behavioural properties of inner speech in healthy participants and patients with aphasia. For this purpose, questionnaires testing inner speech abilities were developed. This chapter describes these questionnaires and their standardisation process. The questionnaires tested participants' ability to detect rhymes, homophones and the location of vocal stress in a word, using inner speech alone. Reading abilities were recorded as well, serving as a measure for overt speech production. The cognitive-linguistic characteristics of inner speech, including its relation to overt speech, were explored in the study.

The following hypotheses were accordingly tested:

1. Based on previous studies, as well as on models emphasising the importance of feedback in speech production, it was hypothesised that participants' performance on the inner speech task will be poorer than that on the overt speech tasks.
2. Comparing performance on tasks involving words with those involving non-words, it was expected that participants will perform better on those tasks involving words due to the lexicality effect.
3. Performance on all tasks was expected to correlate with levels of previous education but not with age.
4. Lastly, it was hypothesised that on the lexical stress questionnaire, participants will show a word length effect. If found, this could be attributed to a load on working memory.

2. Methods and Materials

2.1 Participants

Two lists of the rhyme and homophone judgements were completed by 63 participants (28M/35F; age range: 21-71, mean age: 41.7±20.1; mean number of years of education: 15.6±2.7). Another list of non-word homophones was completed by 97 participants (41M/53F/3data not recorded; age range: 21-72, mean age: 37.1±17.6). Information regarding years of education was recorded only for 63 of these participants. These are the same 63 participants who completed the first two lists of the rhyme and homophone judgement tasks. 28 additional volunteers performed the lexical stress questionnaire (16M/12F; age range: 21-63, mean age: 32.9±10.4). No information regarding education level was recorded for these participants. All volunteers had no previous history of stroke and no history of neurological, psychiatric or language disorders. All volunteers were native speakers of British English. The study was approved by the Cambridge Research Ethics Committee and all participants read an information sheet and gave written consent.

2.2 Materials

Three tests were constructed for assessing inner speech. Participants were asked to perform the tasks 'without pronouncing the words aloud, and without moving any part of the mouth, including lips or tongue'. For a list of stimuli see Appendix 3.

2.2.1 Rhyme and Homophone Judgement - Lists 1 and 2

These two tests were adapted from the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA - Kay, Coltheart et al. 1992). In the first task, participants were asked to determine whether two written words rhyme. For example, 'bear' and 'chair' rhyme, while 'food' and 'blood' do not. The test had altogether 60 pairs. In the second task participants had to determine whether two letter strings sound the same. That is, whether the words are homophones. This test had 40 pairs of real words and 20 pairs of non-words. For example, 'might' and 'mite' are homophone words, while 'ear' and 'oar' are not. Looking at non-words, 'zole' and 'zoal' is a pair of non-word homophones, while 'hane' and 'hine' is a pair of non-words which are not homophones. In the rhyme judgement task, half of the rhyming pairs and half of the non-rhyming pairs had orthographically similar ending (e.g town-gown), while the other half had orthographically dissimilar ending (e.g. chair-bear). This way the test could not be successfully solved based on orthography alone, ensuring that the participants had to use their 'inner speech' to solve the task. The words in all three tasks were randomly assigned to one of two lists.

2.2.2. Non-Word Homophone Judgement - List 3

Since the number of pairs of non-words adapted from the PALPA was small, 40 additional pairs of non-words were developed (see list 3, Appendix 3). These non-words were created by replacing consonants in existing English words. This was done to ensure that the non-words followed English spelling rules. As in the previous tests, this test could not be successfully solved based on orthography alone.

2.2.3 Lexical Stress

The third test examined access to lexical stress in inner speech. Participants were asked to indicate where the stress was in each of 78 different words. The number of items, numbers of syllables and stress location are presented in table 2.1.

Table 2.1. Items in the lexical stress questionnaire.

Number of syllables	Stressed syllable	Number of items	Total number of items
2	1st	16	29
	2nd	13	
3	1st	11	27
	2nd	9	
	3rd	7	
4	1st	4	20
	2nd	9	
	3rd	7	
5	3rd	2	2

Stress was defined to participants as 'some part of a word which is stronger, louder or longer, than others. The stressed part of the word gets the most emphasis'. In the questionnaire, words appeared twice. On the left, the word was printed as a whole, and participants were advised to read the word on this column first, and make their judgement. On the right hand side the word appeared segmented into syllables, and participants were asked to circle the part of the word that is stressed. See examples below.

2.3 Procedure

2.3.1 Rhyme and Homophone Judgements

Participants first performed the entire task using inner speech alone and immediately afterwards were asked to read all the words aloud. This allowed a record to be made of the natural (mostly regional) variability in the pronunciation of the words. The order of lists was counterbalanced between participants.

Scoring of the inner speech task was based on the judgement given to a word pair, with possible answers being correct or incorrect. Hence, every pair judged incorrectly was scored as one error. For the overt reading task, each word was scored separately, so every word read incorrectly was scored as one error. If the participant happened to read both words of a pair incorrectly, this would be scored as two errors. When a self-correction occurred, the word was scored as correctly read aloud.

2.3.2 Lexical Stress

Participants performed the entire task using inner speech alone. They were first given instructions and the following four examples: 1.) In the sentence: 'You are the main suspect!', the stress in the word 'suspect' is on the first syllable: <u>sus</u>pect. 2.) In the sentence: 'I suspect you!', on the other hand, the stress in the word 'suspect' is on the last syllable: sus<u>pect</u>. 3.) In the word 'kingdom', the stress is on the first syllable: <u>king</u>dom. 4.) In the word 'Japan', the stress is on the last syllable: Ja<u>pan</u>. More examples were given if needed.

3. Results

3.1 Rhyme and Homophone Judgement – Lists 1 and 2

First, an item analysis was performed, aimed at excluding any items on which performance was not significantly above chance level. Since all judgements were yes/no, the threshold for exclusion was calculated based on the binomial distribution. Items were excluded if the success rate was at chance level with $p>0.05$. As a result, 5 items (out of 120) were excluded: 3 of the rhyming task (hush-bush, date-plait, gull-full), 2 of the word homophone task (bury-berry, dual-jewel), and none of the non-word homophone task. Additionally, 'success rates' were defined as the percentage of participants giving the correct answer for each item. Items with success rates lower than 70% were excluded. In lists 1 and 2, no such items were found.

Next, the two lists in each test were compared for level of difficulty. Percentage of errors for each item was calculated and the mean percentage of errors for all items on the list was then calculated to define the level of difficulty for that list. Kolmogorov-Smirnov tests showed that the data were not normally distributed (K-S test, $p<0.01$ for all). Therefore, a Mann-Whitney test was used to compare the two lists in each task, revealing that the two lists did not significantly differ in level of difficulty for any of the tests ($p>0.05$ for all). This allows therefore collapsing the data from the two lists when performed in the same condition.

The frequency of the words was calculated, in order to evaluate whether the lists varied in word frequency. Lemma frequencies for the combined spoken and written form were taken from the Celex (Baayen, Piepenbrock et al. 1993). When a word on the list corresponded to two different lemmas (for example, rush can be both a noun and a verb), the combined frequency was calculated. Overall, the four lists did not differ in word frequency (one-way ANOVA, $F_{(182,3)}=0.61$, $p=0.61$).

One participant (70-year-old/M) performed at chance level on some of the homophone judgement tests (homophone list 1: 42% correct, $p=0.14$, and non-word homophones list 1: 30% correct, $p=0.12$). Therefore, data from this outlier were excluded. Data presented below is from 62 participants. Kolmogorov-Smirnov tests indicated that the data deviated significantly from the normal distribution for all tests (K-S test, $p<0.01$). Non-parametric tests were subsequently used.

Table 2.2 illustrates the main variables describing participants' performances on the different tasks.

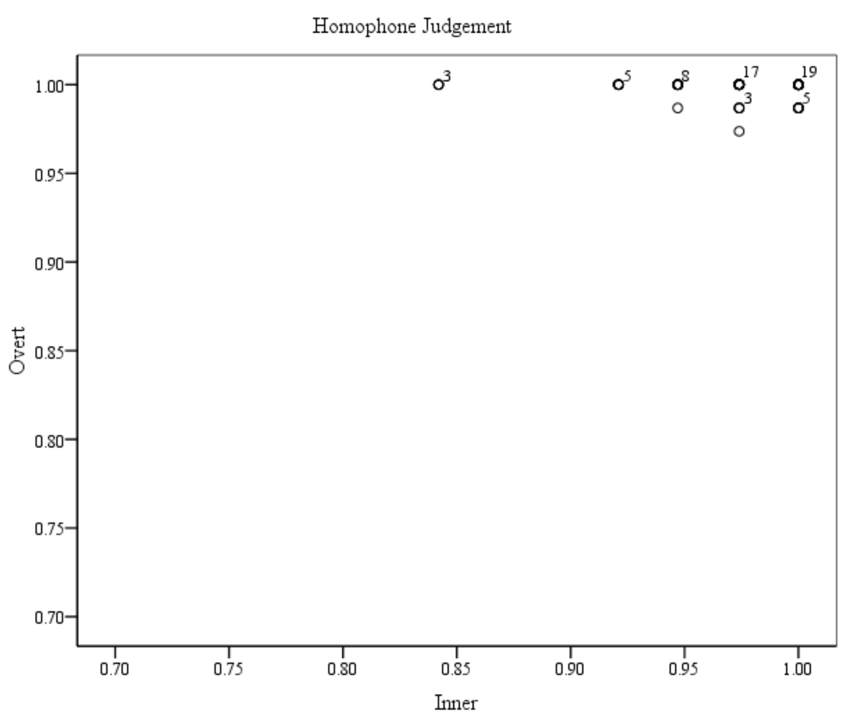
Table 2.2. Participants' performance on the homophone, non-word homophone and rhyme judgement tasks.

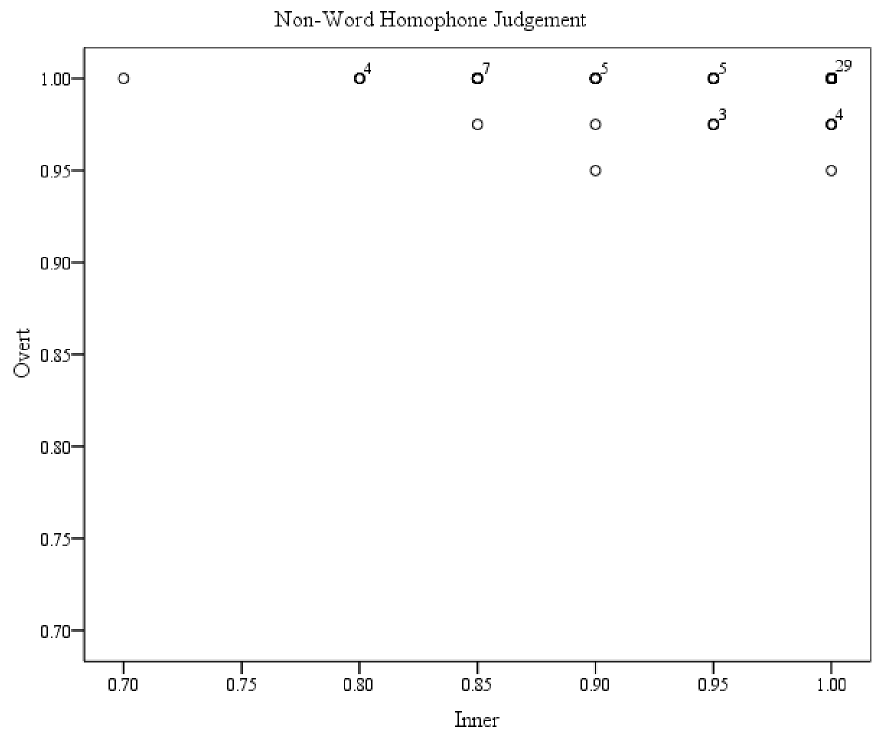
	Condition	Minimum	Maximum	Mean	Std. Deviation
Homophones	Inner	.84	1.00	.97	.04
Non-words homophones	Inner	.70	1.00	.95	.07
Rhymes	Inner	.77	1.00	.96	.04
Homophones	Overt	.97	1.00	.99	.01
Non-words homophones	Overt	.95	1.00	.99	.01
Rhymes	Overt	.97	1.00	.99	.01

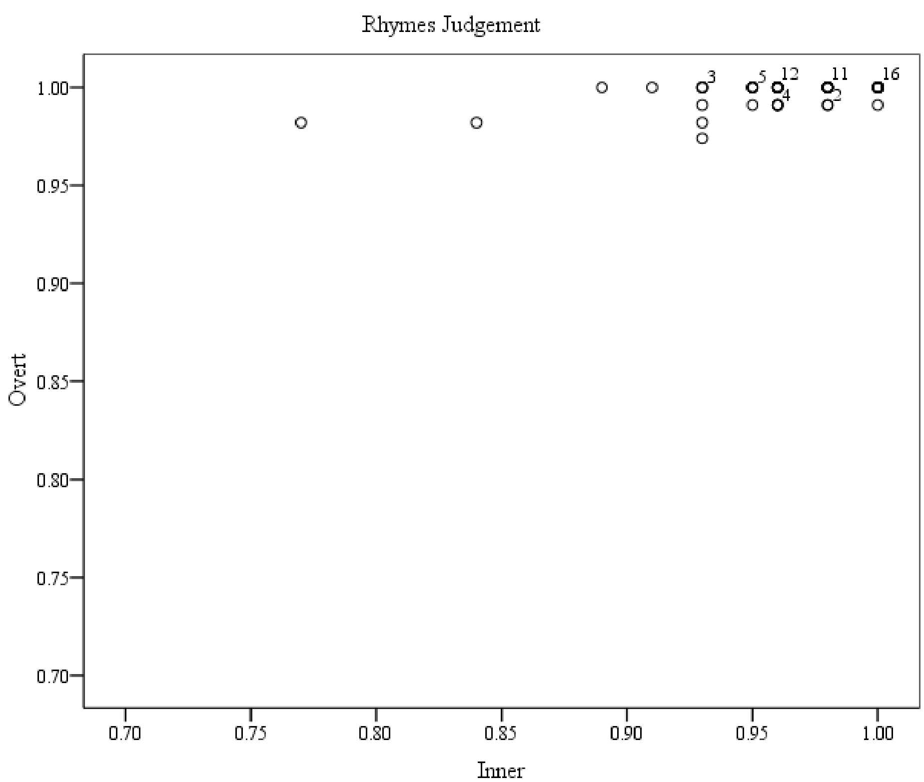
Correlations between the participants' age or amount of formal education, and task performance, were examined using Kendall's Tau (significance level was determined at p<0.004, after correction for multiple comparisons). It was found that age did not correlate with behaviour for any of the tasks (one-tailed Kendall's Tau, p>0.004). However, having more years of formal education positively correlated with performance on the inner speech homophone task. This correlation was significant both for the word condition (Kendall's Tau=0.33, p=0.001) and for the non-word condition (Kendall's Tau=0.27, p=0.004).

Next, the relation between inner speech performance and overt speech performance, in each task, was examined. The correlation between inner speech and overt speech in the homophone task was not significant (words, Kendall's Tau=-0.143, p=0.11; non-words, Kendall's Tau=0.006, p=0.48). However, correlation for the rhyme judgement was significant (Kendall's Tau=0.308, p=0.003). Figure 2.1 shows participants' performance on the inner and overt speech tasks for each of the three tests.

Figure 2.1. Performance on inner and overt speech tasks for each of the three tests: homophone judgement (top), non-word homophone judgement (middle) and rhyme judgement (bottom). The number of cases represented by each dot is notated, when different from 1.



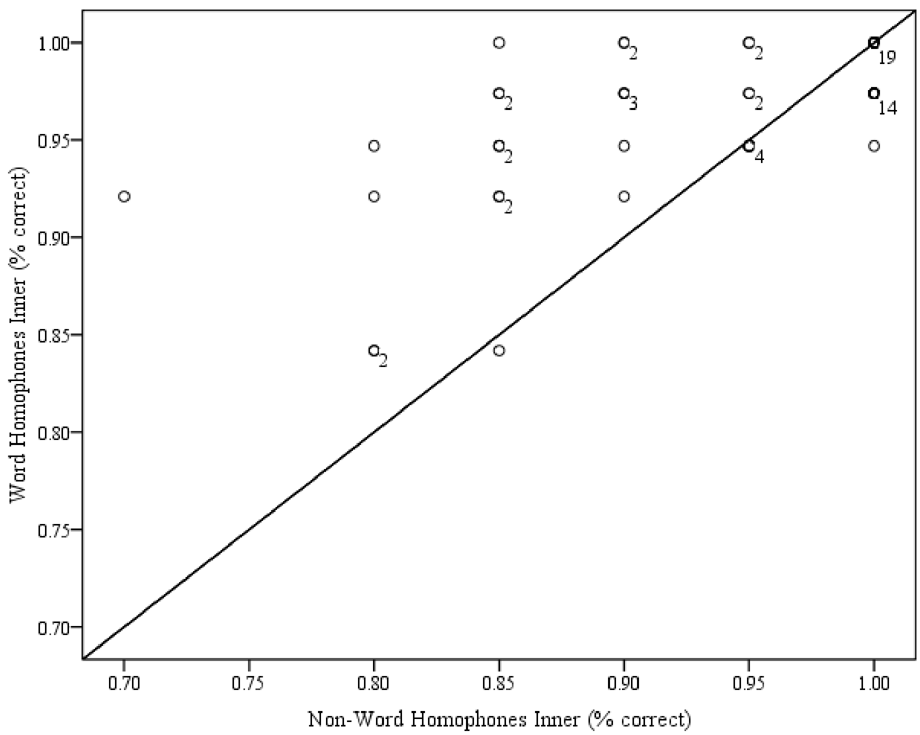

Non-Word Homophone Judgement
Overt
Inner
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.70
0.75
0.80
0.85
0.90
0.95
1.00

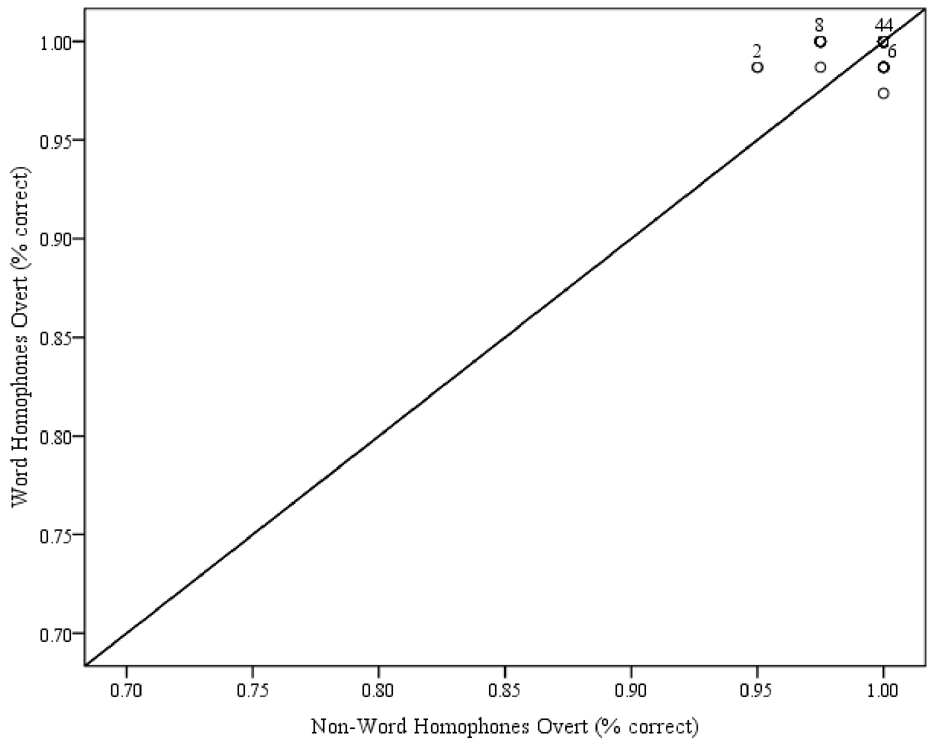

Rhymes Judgement
Overt
Inner
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.70
0.75
0.80
0.85
0.90
0.95
1.00

Looking at the descriptive statistics and figures 2.1 (above) and 2.2 (below), it is clear that the distribution of scores in the inner speech tasks is much wider than that of the overt speech tasks. A one tailed Wilcoxon Signed Ranks Test (comparing dependent non-parametric data) showed that scores for the overt speech tasks were higher than those for the inner speech tasks, in all tests (homophone judgement, $z=-5.323$, $p<0.001$; non-words homophone judgement, $z=-4.547$, $p<0.001$; rhyme judgement, $z=-5.918$, $p<0.001$).

Lastly, performance on the word homophone task and the non-word homophone task was compared. Figure 2.2 shows the relation between the performances on the two tasks.

Figure 2.2. Performance on the word homophone and non-word homophone tasks, using inner speech (top) and overt speech (bottom). The straight line represents equal performance on the two tasks. The number of cases represented by each dot is notated, when different from 1.





In both conditions (inner and overt speech) only one participant showed performance which is substantially better for non-words compared with words. Scores on the task involving words were significantly higher than scores for the task involving non-words both for inner speech (one tailed Wilcoxon Signed Ranks Test, z=-2.583, p=0.005), and for overt speech (one tailed Wilcoxon Signed Ranks Test, z=-1.878, p=.030).

3.2 Non-Word Homophone Judgement – List 3

Lists 1 and 2 of the homophone task had only 40 pairs of non-words. It was therefore desirable to create more such pairs which can then be used to test the aphasia patients. List 3 initially had 40 pairs. An item analysis was first performed, aimed at excluding any items on which performance was not significantly above chance level. This was done in the same way as for lists 1 and 2, and as a result one item was excluded: tousen-towsen (success rate=55%, p=0.53). Further five items were excluded based on low success rates: shink-shynke (success rate=32%), maccep-maxep (40%), ama-amah (57%), qaffal-kaphal (63%) and thirm-theerm (68%). One participant (44-year-old/F) performed at chance level (62% correct, p=0.054) and her data was excluded. After excluding items and participants based on these analyses the standardised list was comprised of 34 non-word pairs and data from 96 participants was included. Overt speech was recorded for 32 of these participants (13M/20F; age range: 21-71; mean age=47.4±20.3; mean number of years of education=16.4±2.9).

Table 2.3 illustrates the main variables describing participants' performances on the different tasks.

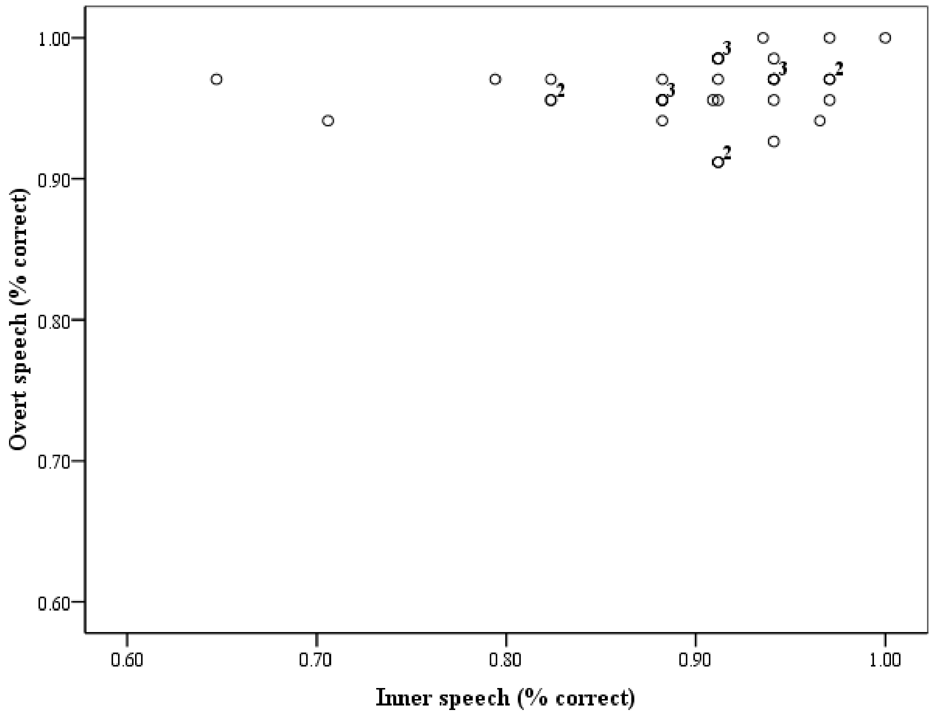
Table 2.3. Participants' performance on the inner and overt tasks, list 3 of the non-word homophone judgement.

Condition	Minimum	Maximum	Mean	Std. Deviation
Inner	.65	1.00	.89	.08
Overt	.91	1.00	.96	.02

Kolmogorov-Smirnov tests indicated that the data deviated significantly from the normal distribution (K-S test, $p<0.01$). Therefore, non-parametric tests were subsequently used. Correlations between age or amount of formal education, and task performance, were examined using Kendall's Tau. It was found that neither age nor amount of education correlated with behaviour for any of the tasks (one-tailed Kendall's Tau, $p>0.05$).

Lastly, the relation between inner speech performance and overt speech performance was examined. It was found that the correlation between inner speech and overt speech did not reach significance ($n=32$, Kendall's Tau=0.21, $p=0.066$). Figure 2.3 shows the relation between inner and overt speech performance on the non-word homophone task, list 3.

Figure 2.3. Performance on the inner and overt speech tasks, non-word homophones, list 3. The number of cases represented by each dot is notated, when different from 1.



From table 2.3 and figure 2.6 it can be seen that similarly to lists 1 and 2, the range of scores and the variance was larger for the inner condition compared to the overt condition. Scores on the task involving overt speech were significantly higher than scores on the task involving inner speech (one tailed Wilcoxon Signed Ranks Test, z=-4.343, p<0.001).

3.3 Lexical Stress

First, an item analysis was performed, aimed at excluding any items on which performance was not significantly above chance level. Chance level was calculated using the chi squared test, taking into account the number of syllables in the word. Therefore, for a word with two syllables, the probability of giving a correct answer by chance is 0.5. For a word with 3 syllables, the probability is 0.33, for a word with 4 syllables - 0.25, and for a word with 5 syllables - 0.2[9]. One item was excluded based on this analysis: 'hotel' (2 syllables, success rate=56%, p=0.11). Further 12 items were excluded based on low success rates. These included 8 items with 3 syllables ('difficult', success rate=68%; 'committee', 68%; 'understand', 64%; 'comprehend',

[9] This method makes some assumptions which will be highlighted in the discussion of this chapter.

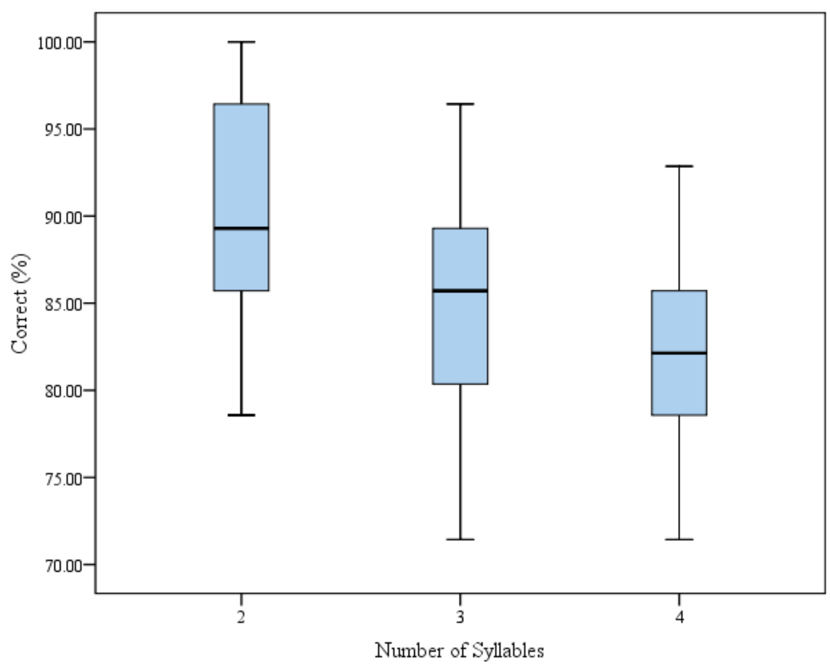
64%; ‘elephant’, 61%; ‘photograph’, 61%; ‘contradict’, 61% and ‘hurricane’, 57%) and 4 items with 4 syllables (‘escalator’, success rate=68%; ‘information’, 68%; ‘elevator’, 55% and ‘economic’, 55%).

Next, it was examined whether any of the participants received a total score not significantly above chance level. The chance level for the entire questionnaire was calculated based on the numbers of syllables in each word and the number of items with each number of syllables. The score not significantly different from chance was found to be 58%. This means that participants who scored 58% correct or lower were excluded. This resulted in the exclusion of data from two participants (22-year-old/M who scored 48% correct and 34-year-old/F who scored 46% correct).

In summary, the data for 65 items (28 two-syllabic words, 19 three-syllabic words, 16 four-syllabic words and 2 five-syllabic words) and 26 participants were included, after performing the standardisation procedure described.

The distribution of items’ success rate did not significantly deviate from the normal distribution (K-S test, $p>0.05$). Therefore, parametric tests were used. The success rates for words with different numbers of syllables were compared. Figure 2.4 shows the success rate for words with different numbers of syllables, across all participants.

Figure 2.4. Success rate for words with different numbers of syllables, lexical stress task.



The group of words with 5 syllables was excluded from this analysis since it only had 2 items. The difference between success rates approached significance (one-way ANOVA, $F_{60,2}$=2.97, p=0.059). This suggests that shorter words received more correct answers. Planned comparisons showed that while the words with 2 and 3 syllables differed significantly in their success rates (independent sample t-test, t=3.07 p=0.004), words with 3 and 4 syllables did not significantly differ from each other in their success rates (independent sample t-test, t=1.82 p=0.078).

Participants' scores on the lexical stress questionnaire ranged from 0.6 to 1.0, with a mean score of 0.896±0.124. The data deviated significantly from the normal distribution (K-S test, $p<0.05$), therefore, Kendall's Tau was used to test for correlation between age and performance on the task. It was found that the two variables did not correlate significantly (Kendall's Tau=0.016 p=0.911).

4. Discussion

4.1 Standardisation of Tests

This study aimed to standardise questionnaires specifically designed to examine inner speech. Healthy native speakers of British English were tested and their performance was recorded. These scores can be now used to assess the performance of patients with language impairments. Many cognitive neuropsychology tests use a cut-off score (usually z=-2) to define impaired and intact performance. This was done in tests such as the Mini Mental State Examination (Folstein, Folstein et al. 1975), Hayling and Brixton Tests (Burgess and Shallice 1997), Wechsler Adult Intelligence Scale (Wechsler 1955) and the Wisconsin Card Sorting Test (Berg 1948), to name but a few. Cut-off scores should be calculated from a sample of healthy volunteers which is matched to the patient population on relevant variables. In this study, it was found that level of education correlated with performance on the inner speech tasks. Therefore, a control group should be matched with the patient group on this variable. It is possible to select some of the healthy volunteers who participated in this study, to create a group matched to a patient group on the desired variables, and calculate the relevant cut-off score based on this sub-population.

Level of education, but not age, correlated with performance on some of the tests. The relation between language abilities and formal education is well documented. For example, Zanini et al. (2005) found that level of education was correlated with verbal communication ability, in a sample of Italian healthy adults. Similar results were found in a group of Mexican adults (Ardila, Ostrosky-Solis et al. 2000) and African and white Americans (Marcopulos, McLain et al. 1997; Manly, Jacobs et al. 1998). Apart from the study by Meijer et al. (2008), these studies did not find a correlation between age and language performance. Manly (1999) suggests that it is the quality, rather than the amount of education that influences language ability. However, in the current study it was impossible to objectively assess the quality of education participants received. The influence of education level on language ability can be complex and it interacts with other related variables such as socio-economic status (Dotson, Kitner-Triolo et al. 2009), racial background (Marcopulos, McLain et al. 1997; Dotson, Kitner-Triolo et al. 2009), age (Ardila, Ostrosky-Solis et al. 2000) or literacy (Manly, Jacobs et al. 1999). It was found that correlation between performance on the rhyme judgement and level of education was not significant. This could be a result of the influence of variables which were not measured, as the ones mentioned above (for example, literacy or socio-economic status). However, there is an important difference between the rhyme and homophones judgement. While rhyme judgement requires working memory abilities, studies of homophone judgement show that performance on this task is not affected by articulation suppression (reviewed in Howard and Franklin 1990), implying that homophone judgement task does not require working memory. One can therefore hypothesis that rhyme judgement was influenced by working memory abilities, over and above potential correlations with other variables such as education level.

In summary, the data collected can be used to evaluate inner speech abilities of patients with language impairments. Clearly, testing a larger sample of healthy volunteers, and measuring other independent variables such as socio-economic status or literacy, could make the data more reliable and applicable for wider populations. Moreover, regional variations in accent might be important in the standardisation of the tests. This was not formally evaluated in this study. However, looking at the rhyme and homophone judgements, one might notice that 3 (hush-bush, gull-full and bury-berry) out of the 5 pairs that were excluded, have clear regional variability. For example, while an English speaker from the Cambridge area might pronounce ‘hush’ and ‘bush’, or ‘gull’ and ‘full’ as non-rhymes, an English speaker from Yorkshire is more likely to pronounce these words as rhyming pairs. The opposite is true for the pair ‘burry’ and ‘berry’ which is usually pronounced as a homophone pair by residents of southern England but not by people from the north of England. In summary, the influence of variability in accents on the judgements given should be systematically explored in the future.

4.2 Relation between Inner and Overt Speech

Performance on the inner speech tasks was poorer than that of overt speech tasks in all cases. Two explanations, briefly mentioned in the introduction, can account for this finding. Firstly, many models of language production emphasise the importance of feedback, with three types of feedback proposed: inner speech or the articulatory buffer (Levelt 1983; Levelt 1989; Levelt 1993; Postma 2000), motor feedback (Guenther 2005), and feedback from overt speech via the comprehension system (Guenther 2005). Inner speech is supported only by the first type of feedback, while overt speech is supported by all three types of feedback. Hence, poorer performance on inner speech tasks, compared to overt speech tasks, can be attributed to lack of some of the feedback streams during inner speech production. And indeed, as mentioned above, healthy participants vocalise more as the difficulty of reading material increases, and have lower comprehension scores when prevented from vocalising or even subvocalising (Hardyck and Petrinov 1970).

Another explanation for the difference in scores between inner and overt speech is that inner speech is ‘phonetically impoverished’ in comparison to overt speech (Oppenheim and Dell 2008). If this is true, than the consequence might be that phonological tasks, such as the ones used in this study, might be better performed when overt speech is used than when inner speech is used. Regarding the rhyme judgement, one might argue that poorer performance can be attributed to working memory load or difficulty. This might be a relevant explanation when examining patients with brain damage, who might have working memory deficits. However, there is no reason to suggest that the healthy participants who participated in this study had poor working memory. Normal working memory, in the absence of any distraction, should be sufficient for a simple rhyme judgement task.

4.3 Differences between Words and Non-Words

A lexicality effect was found in all cases: performance on tasks using words was significantly better than performance on tasks using non-words. These findings replicate previous results showing that performance on tasks using non-words is worse than performance on word tasks, as seen in the performance of a patient and a group of healthy controls in Baddeley and Wilson's study (1985). Despite genuine differences between various models of speech production, phonological processing, reading aloud, etc. (see for example Paap and Noel 1991; Coltheart, Curtis et al. 1993; Martin, Saffran et al. 1996; Plaut, McClelland et al. 1996; Dell, Schwartz et al. 1997), all models incorporate one or more sources of lexical/semantic support for phonology, with the inevitable consequence that, unless performance on non-words is at ceiling, scores on word tasks will always exceed those on non-words. In general, these models can easily account for these results since words receive activation from two sources: semantic and phonological, while non-words receive only one type of activation: phonological (Seidenberg and McClelland 1989; Paap and Noel 1991; Coltheart, Curtis et al. 1993; Martin 2003). Moreover, the connections between the phonological units of a word are well practiced and are therefore stronger than that of non-words (Acheson and MacDonald 2009). Having these extra sources of activation increases the chance of choosing the correct phonological unit and therefore reduces errors (Wilshire 2008). This semantic influence on speech production is supported by these results. Semantic influences on inner speech are not previously documented, but, as hypothesised, were found in this study.

4.4 Lexical Stress

It was found that in the lexical stress questionnaire, performance on words with 2 syllables, was better than performance on words with 3 syllables (the difference between words with 3 and 4 syllables was close to significant but did not reach the critical cut off point of $p=0.05$). This can be explained as a working memory influence: it is more difficult to hold longer words in working memory and analyse them, compared to shorter words. However, as words get longer, the probability of giving a correct answer by chance, without any previous knowledge, decreases. This confound could not be dealt with in this study. It is therefore impossible to reach a confident conclusion regarding the relation between word length and the ability to determine its lexical stress in inner speech.

In order to define success in the lexical stress questionnaire, the chance level was calculated according to the chi squared distribution. This method was used because of the lack of a better method. However, it makes a problematic assumption.

Specifically, by determining the chance level for stress assignment in a 4-syllabic word as ¼, for example, one makes the assumption that speakers have no knowledge of language patterns. However, this is not completely true. It is well-known that stress assignment in English obeys specific rules. The most frequent pattern is to stress the first syllable of a word with 2 syllables (Cutler and Norris 1988). This general rule suffices to correctly determine the stress of 83% of all disyllabic English words in the CELEX database (Rastle and Coltheart 2000). More importantly, it was found that speakers are sensitive to these rules. This was seen in two experiments; one examining stress assignment in non-word reading, and the other, analysing reading latencies and errors in stress assignment when reading aloud real words (Rastle and Coltheart 2000). Hence, the probability of correctly determining stress location by chance might have to take into consideration the regularity of the stress patterns in each word in the questionnaire. This is a complicated task because the definition of regularity is by no mean under consensus. For example, while some scholars argue that lexical information influence stress assignment (for example whether the word is a verb or a noun, Kelly and Bock 1988), others suggest that lexical information is not essential (Rastle and Coltheart 2000).

The next question that arises is then: when do speakers access the information about a specific word's stress? Rastle and Coltheart (2000) developed one of the only reading models which aims at explaining how stress is determined by readers. They argue that stress is defined by rules, rather than being an information which is attached to each and every word. However, they suggest that rules can be derived from orthographic, as well as phonological, information. If rules are derived from orthographic information, then theoretically, speakers can determine stress before accessing the word form, i.e. without inner speech. If rules are derived from phonology, on the other hand, then an impoverished form of inner speech (without stress location) is a prerequisite for stress assignment. The authors do not opt for one of the options over the other.

Levelt (1999), on the other hand, argues that the word stress is specified during the level of morphophonological encoding. This process involves the retrieval of a few types of information; among them are the metrical shape of the word, specifying the stress pattern[10], and the segmental makeup, specifying the phonemes

[10] Levelt argues that stress is only specified when it is different from the language default.

used to build the word. The various types of information are then combined to create the phonological word. This implies then, that stress pattern is available to the speaker before inner speech is produced.

There is some evidence suggesting that people indeed have access to the metrical shape of the word in the absence of access to the word form itself. Barton (1971) examined 16 patients with aphasia, who were able to determine the number of syllables of unnamed words, above chance. Similarly, it has been found that conduction aphasics were superior to Wernicke's and anomic aphasics, in their ability to determine the syllabic length of words they could not retrieve (Goodglass, Kaplan et al. 1976). In their seminal study of tip-of-the-tongue states (TOT), Brown and McNeill (1966) asked students to guess the number of syllables of words they could not retrieve. Participants were correct in 60% of their guesses in the main study and 47% in a pilot investigation. These studies suggest that speakers have information regarding the metrical shape of the word even without retrieving the word form. However, there are no studies to date showing that speakers can access stress without accessing the word form. Examining TOT states, Rubin (1975) showed that when participants retrieve words which are related to the target word, these are likely to have stress patterns similar to the target word. This data can be interpreted as providing evidence that speakers have knowledge of the stress of words they cannot name. However, Rubin acknowledges that this result can also emerge from the fact that the related words are phonologically similar to the target words, and therefore have a similar stress pattern.

In summary, the literature does not give clear answers as to how and when speakers access lexical stress, and whether stress can be accessed without accessing the word form. Answering this question seems essential if one is to use the lexical stress questionnaire to assess inner speech. It is suggested that reaction times (RT) might shed light on which strategy people are using, with fast RTs suggesting no inner speech, and long RTs suggesting the use of inner speech. At the moment, however, this question remained unsolved. Another unsolved problem is the question of how to choose a meaningful baseline for the calculation of success rates, as explained above. Taken together, it seems that although speakers testify that they use inner speech when performing the lexical stress task, it remains problematic to use it as a tool for assessing inner speech.

4.5 Conclusions and Future Directions

The main aim of this study was to develop questionnaires which can assess inner speech abilities. This aim was achieved and three main types of questionnaires were developed: rhyme judgement of words, homophone judgement of words and homophone judgement of non-words. A lexical stress questionnaire proved to be problematic as a tool for assessing inner speech. It was found that overall, participants' performance on inner speech tasks is worse than performance on overt speech tasks, and that the lexicality effect can be found in both inner and overt speech tasks. Lastly, it was found that amount of formal education, but not age, was correlated with performance on some of the inner speech tasks.

The next chapter describes a study in which inner and overt speech abilities were studied in a population of patients with post-stroke aphasia.

Chapter 3

Behavioural Investigation of Inner Speech in Aphasia

"The auditory type," says M. A. Binet, "appears to be rarer than the visual". Persons of this type imagine what they think of in the language of sound. In order to remember a lesson they impress upon their mind, not the look of the page, but the sound of tile words… Mozart, for example, noting from memory the Miserere of the Sistine Chapel after two hearings; the deaf Beethoven, composing and inwardly repeating his enormous symphonies. On the other hand, the man of auditory type… is exposed to serious dangers; for if he loses his auditory images, he is without resource and breaks down completely.

William James, Principles of Psychology (1890)

1. Introduction

The nature of inner speech in aphasia occupied neurologists, psychologists and aphasiologists since the 19th century. Brown (2009) reviews the developments in the definition of inner speech among the early famous aphasiologists. Wernicke (1848-1905) termed the word '*Wordbegriff*', or 'word-concept' and defined it as the combination of acoustic and motor speech imagery. He attributed the control of inner speech to the arcuate fasciculus which connects the posterior auditory areas to the anterior motor language areas. Dejerine (1849–1917) talked about the 'true inner voice' or '*une veritable voix interieure',* which consisted of a combination of auditory and motor imagery. He argued that true aphasia involves the loss of inner speech. Later, Goldstein (1878-1965) argued that inner speech is the process connecting word finding and articulation. He attributed phonemic paraphasias to a disruption of inner speech (Brown 2009). Luria (1902–1977) described a patient who could understand simple sentences and perform basic arithmetic when he was free to speak out loud. However, when his lips and tongue were clamped, his comprehension and arithmetic performance declined considerably. Luria concluded that the patient had impaired inner speech but preserved overt speech. When prevented from producing overt speech he could not rely on his inner speech and therefore his performance broke down (Luria 1963). Luria argued that inner speech is the process connecting thought and articulation, and claimed that the formulation of an utterance cannot be accomplished without inner speech. He maintained that inner speech is not simply overt speech without the articulation component, and discussed the involvement of impairment of inner speech in dynamic aphasia (Luria 1964; Luria 1976). Weigel (1901-1979) followed this line of research and in a study published in 1964 he prevented patients with aphasia from using their articulatory system during listening, by asking them to blow into a tube, as hard as they could. Patients could still comprehend sentences and repeat words (translated in De Bleser and Marshall 2005). It is therefore clear that defining the concept of inner speech and understanding its role in speech production engaged the early aphasiologists to a large extent. However, it seems that more recent investigators have lost some of this extended interest in inner speech. As a result, contemporary aphasia literature has a lack of experimental studies as well as theoretical work investigating the nature of inner speech and its characteristics in aphasia. Below is a review of experimental literature investigating inner speech in post-stroke aphasia, published in the second half of the 20th century.

Aphasic patients often complain that there is poor correspondence between the words they think or intend to say, and the words they are able to produce out loud (Marshall, Neuburger et al. 1994). Is there evidence to show that such a claim is indeed true? Patients with impaired speech due to motor dysfunctions (i.e. those suffering from dysarthria, apraxia or aphemia) may still have the ability to produce inner speech and to use it successfully to complete language tasks (Nebes 1975; Baddeley and Wilson 1985). Looking at patients without severe motor speech deficits, Pratap (1987) found normal speech evoked potentials during silent repetition for an anomic patient with a left hemispheric infarct. However, it is not reported whether the patient could repeat the words aloud as well. Feinberg and colleagues (Feinberg, Gonzalez Rothi et al. 1986) tested five patients with conduction aphasia who were unable to read words aloud. Four of the five demonstrated spared performance on inner speech tasks such as judgement of word length, and judgement of whether pairs of words were homophones or rhymes, all using pictures. Marshall et al. (1985) presented a case study of a patient who had severe auditory comprehension deficits and impairment in speech production. Despite this, she often corrected her own errors and was relatively successful on various inner speech tasks, including rhyme and homophone judgement and phoneme monitoring in reading. However, when asked to determine syllable number using inner speech she resorted to overt speech. These case studies suggest that there is dissociation between inner and overt speech, although it should be noted that the authors of all did not assess, or at least did not report, presence or absence of speech apraxia. It is therefore difficult to establish what component of the speech impairment was attributable to motor speech deficits. Lastly, Ross (1983) tested patients with either posterior or anterior left hemispheric brain damage, on their performance on a silent reading task. In the study, participants were asked to cross out a specific letter whenever it occurred. Normal readers easily detect a letter when it has a regular pronunciation, but tend to ignore it when it has an irregular pronunciation (Corcoran 1966, for details of this study see chapter 1). Ross found that as a group, the aphasia patients used inner speech while reading a paragraph for comprehension, since their error rate was no different from patients with right hemispheric brain damage (Ross 1983). However, the author notes that there was considerable variability within the patient group, with two patients having a 70% error rate, which is high above the normal rate. However, this variability is not discussed.

Another source of evidence regarding inner speech in aphasia comes from the examination of error correction in patients. Levelt (1989) proposed that errors can be corrected through two pathways: a pre-articulatory loop (i.e. inner speech, discussed before) responsible for detecting errors before they are produced, and a post-articulatory monitoring mechanism responsible for detecting errors after they are produced through the auditory comprehension system (Levelt 1989). Schlenck et al. (1987) found a low occurrence of repairs (self-corrections after overt errors) among aphasic patients, which is indicative of malfunction of the post-articulatory monitor. By contrast, the number of 'pre-pairs' (self-corrections without overt error which were defined as 'searching behavior aiming at a following utterance' ibid, p. 232) was high, which indicates a better preserved pre-articulatory monitoring mechanism. However, the interpretation of the results is based on the problematic assumption that all overt repairs are post-articulatory (Oomen, Postma et al. 2001). Oomen et al. (2001) tried to distinguish between post- and pre-articulatory corrections of speech errors, by using white noise. In the presence of white noise people cannot hear themselves speak, and therefore can only use the pre-articulatory loop for self-repair. Control participants made more corrections in the absence of white noise (i.e. when able to rely on both pre- and post-articulatory mechanisms) than in the presence of white noise, indicating that they can benefit from both repair mechanisms. Patients with Broca's aphasia, on the other hand, made the same number of corrections in the two conditions. Moreover, the patient group made fewer corrections than the normal group in the absence of noise. In line with Schlenck et al. (1987), these data suggest that Broca's aphasics have an impaired post-articulatory monitor. Nevertheless, they can sometimes use their pre-articulatory monitor mechanism to correct their speech errors.

These studies not only demonstrate that inner speech might be at least partially preserved in many cases of aphasia (Feinberg, Gonzalez Rothi et al. 1986; Pratap 1987; Schlenck, Huber et al. 1987; Oomen, Postma et al. 2001), but also that some patients learn to rely on inner speech more than on overt speech comprehension for correcting their speech errors (Oomen, Postma et al. 2001). Oomen et al. (2001) went further in suggesting that '…speech fluency of Broca's aphasics might be increased when they are trained in therapy… by means of the auditory loop' (ibid. p. 640). It is important to note, however, that positive findings regarding inner speech in aphasia are not universal. Three studies have reported complete absence of inner speech: one in a patient with conduction aphasia (Levine, Calvanio et al. 1982), one in a patient with severe aphasia due to a posterior damage (Martin and Caramazza 1982) and the last, in a patient with progressive anomia due to left temporal lobe atrophy (Snowden and Neary 2003).

In summary, data regarding inner speech in post-stroke aphasia are scarce, incomplete and inconclusive. Inner speech is not routinely compared to overt speech abilities, nor is inner speech ability a part of the regular diagnostic procedure employed by most speech and language therapists. Most studies that have examined inner speech have tended to treat it as an existing or absent ability, while in reality patients might show graded levels between these two extremes. It is also clear that most previous studies did not take motor speech deficits into consideration when examining inner and overt speech.

The current study was aimed at furthering our understanding of the competence and role of inner speech in aphasic language processing in the context of current language models. It asked the following questions:

1. Do stroke patients with aphasia typically have impairments of inner speech? Based on previous studies, it was hypothesised that some patients have preserved inner speech while the others do not.

2. Is there any behavioural dissociation between inner and overt speech abilities? It was hypothesised that some patients will show the pattern seen by healthy volunteers: performance on overt speech tasks is better than performance on inner speech tasks. However, language models predict that, after brain damage, the opposite pattern will occur as well.
3. Do patients show a lexicality effect in inner and overt speech? It was hypothesised, that like the healthy controls, patients will show lexicality effect in both inner and overt speech.
4. Do phonological judgements made by patients rely on inner speech or overt speech when the two differ? Previous studies show that patients rely on inner speech for error monitoring. However, when inner speech is impaired, patients might rely solely on overt speech. It was therefore hypothesised that both patterns should occur in the patient population.

To investigate these questions, rhyme and homophone judgement tasks were used to assess inner speech and a reading aloud task to assess overt speech, with words and non-words used in both cases. Both the rhyme and the homophone judgement tasks require the subject to retrieve the phonological form of the word, translate it to an internal auditory representation, and examine its properties before making an overt decision. However, the tasks differ in that rhyme judgement appears to require working memory abilities, as well as inner speech abilities, while homophone judgement does not require working memory. It was mentioned in the introduction to this thesis that the phonological loop, the verbal part of the working memory system, is made of two sub-components. The first, the phonological store, is where phonological information is held. The second, the articulatory control process, is an active processor which allows refreshing the memory trace of the information inside the phonological store by using subvocal rehearsal. Visual input accesses the phonological store through the articulatory control process, while auditory input can access the phonological store directly. Articulating irrelevant material will disrupt the normal function of the articulatory control process but not of the phonological store. This process is known as *articulatory suppression*. It has been shown that when participants are asked to recite irrelevant material aloud, their performance on rhyme judgement, but not on homophone judgement, declines: they make more errors and their responses are slower. This shows that rhyme judgement, but not homophone judgement, is affected by articulatory suppression. This data has been taken by many as evidence to support the idea that rhyme judgement requires the availability of the phonological loop, while homophone judgement does not (reviewed in Howard and Franklin 1990).

To verify that poor performance on the rhyme judgement task cannot be solely attributed to a working memory deficit, patients' performance on other tasks, which require working memory, was also examined. This included sentence repetition (Saffran and Marin 1975; Caramazza, Basili et al. 1981) and sentence comprehension (Vallar and Baddeley 1984). If patients perform well on these tasks, it seems unlikely that their working memory is impaired to any significant extent. If they then perform poorly on the rhyme judgement, then this poor performance can be attributed to impairments of inner speech rather than impairments of working memory.

2. Methods and Materials

2.1 Participants

2.1.1 Patients with Aphasia

29 individuals with chronic (at least 6 months post-stroke) aphasia were recruited (20M/9F; age range: 21-81 years, mean age: 64.6±13.4 years; mean number of years of education: 12.7±3.3; mean time since last stroke: 28.8±25.3 months). All patients had left Middle Cerebral Artery (MCA) territory stroke, were above 18 years of age, and were native speakers of English, with all but one being native speakers of British English. 23 patients were right handed, 4 were left handed, and 2 were defined as ambidextrous. Handedness was assessed using the Edinburgh Handedness Inventory (Oldfield 1971). The diagnosis of aphasia was based on the convergence of clinical consensus and the results of a standardised aphasia examination - the Comprehensive Aphasia Test (CAT - Swinburn, Porter et al. 2004). Patients had impaired production of speech but relatively preserved comprehension (to a level allowing them to give informed consent to the study and understand the behavioural tasks). Patients had no history of other neurological or psychiatric disorders and no major cognitive impairment other than the language deficit which might prevent them from understanding and performing the different tests. Table 3.1 presents additional

demographic and clinical information. Two patients, whose demographic and clinical data are not presented underneath, were unable to complete the behavioural tests successfully because of their impairment in language comprehension, and were therefore excluded from further analysis. Appendix 4 documents the clinical case histories of each patient together with brain images depicting the lesion site, when available.

2.1.2 Healthy Controls

The group of normal controls was a sub-group of the healthy controls group who participated in the standardisation of the questionnaires (chapter 2). This sub-group matched the patient group on age and amount of formal education, and was comprised of 27 healthy volunteers (12M/15F; age range: 44-72, mean age: 62.2±8.2 years; mean number of years of education: 13.9±2.6). All volunteers had no previous history of stroke and no history of neurological, psychiatric or language disorders. All volunteers were native speakers of British English. The study was approved by the Cambridge Research Ethics Committee and all participants read an information sheet and gave written consent.

Table 3.1. Patients' demographic and clinical information

Patient		Age	Sex	Edu[1]	Hand[2]	Type of Stroke	Test completed (inner speech battery)		
							Rhymes	Word homophones	Non-word homophones
1	AE	69	m	BA	R	ischaemic	+	+	-
2	AP	73	m	age 11	L	ischaemic	+	+	-
3	AS	48	f	MA	R	ischaemic	+	+	+
4	BA	62	m	MA	R	haemorrhagic	+	+	+
5	BBS	78	m	age 17	A (-0.1)	ischaemic	-	+	+
6	DB	78	f	age 14	R	ischaemic	+	+	+
7	DH	78	m	age 14	L	ischaemic	+	+	-
8	GH	42	f	HS	R	ischaemic	+	+	+
9	HD	81	m	age 16	R	ischaemic	+	+	-
10	IB	62	m	PhD	R	ischaemic	+	+	+
11	IC	61	m	age 16	R	ischaemic	+	+	-
12	JFI	65	f	age 15	R	haemorrhagic	+	+	+
13	JFO	71	m	age 15	R	ischaemic	+	+	+
14	JHH	71	M	Age 15	R	ischaemic	+	+	+
15	JS	79	m	HS	L	ischaemic	+	+	+
16	LM	49	f	BA	R	ischaemic	+	+	+
17	MM	70	m	HS	A (0.4)	ischaemic	+	+	+
18	NET	70	m	age 16	R	ischaemic	+	+	+
19	PT	65	m	HS	R	ischaemic	+	+	+
20	RB	53	f	BA	R	ischaemic	+	+	+
21	TH	75	f	BA	L	haemorrhagic	+	+	+
22	WS	51	F	Age 12	R	ischaemic	-	+	-
23	CB	69	m	HS	R	ischaemic	+	+	+
24	DR	73	m	Age 16	R	ischaemic	+	+	+
25	EB	21	f	HS	R	ischaemic	+	+	-
26	CR	60	m	Age 16	R	ischaemic	+	+	+
27	AD	66	m	HS	R	ischaemic	+	+	+

[1] Education level: HS=finished high school, BA=Bachelor's degree, MA=Masters degree, PhD=Doctorate degree.

2 L=Left, R=Right, A=Ambidextrous. In brackets: the score achieved on the Edinburgh Handedness Inventory for ambidextrous participants, where (-1)=strongly left handed, (1)=strongly right handed and (0)=completely ambidextrous.

2.2 Materials

The following tests were administered to the volunteers: the Raven Matrices (for measuring non-verbal IQ, Raven 1938), the inner speech test battery, and a personal questionnaire (assessing knowledge of other languages and general medical background). Patients were tested on the Raven Matrices (Raven 1938), the Comprehensive Aphasia Test (Swinburn, Porter et al. 2004) and the Apraxia Battery for Adults (ABA - Dabul 1979). Patients were also given a set of cognitive tests, including the Brixton Test of executive functions (Burgess and Shallice 1997), the Rey - Osterreith Complex Figure Test (Meyers and Meyers 1995) and parts of the Addenbrooke's Cognitive Examination - Revised (ACE-R), testing visual-spatial abilities (Mathuranath, Nestor et al. 2000). These are summarised in table 3.2.

Table 3.2. Behavioural tests administered to the healthy volunteers and to the patients with aphasia.

Healthy volunteers	**Patients with aphasia**	**Attributes assessed**	**Time to Administer**
Consent form	Consent form		10-15 minutes
Safety form	Safety form	Suitability for MRI	
Edinburgh Handedness Inventory	Edinburgh Handedness Inventory	Handedness	
Raven Matrices	Raven Matrices	Non-verbal IQ	20 minutes
	Comprehensive Aphasia Test	Language ability	$1^1/_2$ –3 hours
	Apraxia Battery for Adults	Symptoms of apraxia	10 minutes
Personal questionnaire	Personal questionnaire	Knowledge of foreign languages; medical background of healthy volunteers	5 minutes
Inner speech test battery - Rhymes - Homophones	Inner speech test battery - Rhymes - Homophones	Inner and overt speech abilities	20 minutes
	Rey–Osterrieth Figure	Short term memory	20 minutes
	Brixton Test	Executive functions	10 minutes
	ACE-R	Visuo-spatial abilities	10 minutes

For details about the specific tests see Appendix 5.

2.3 Procedure

For patients, tests were administered in 2-3 sessions, depending on the patient's ability. Test sessions took place either at Addenbrooke's hospital or at the patient's home, according to the patient's preference. All healthy volunteers were tested in one session each, in Addenbrooke's hospital.

For each of the tests in the inner speech battery, items were randomly divided into two lists (see Appendix 3). Patients performed the task on half of the items (one list) using inner speech and half (second list) using overt speech. In both cases, the patient would first read both words in the pair (either internally or overtly, depending on the condition) and then give his/her judgement for the pair. This way the difference between inner and overt speech can be quantified. It also allowed estimating whether patients rely on inner speech or on overt speech when making the rhyme and homophone judgements during overt production (for further details see results below). The two conditions were completed separately and successively, and the order of conditions was randomised between patients. Lists were randomised between patients so half of the patients judged list 1 using inner speech and list 2 using overt speech, and the other half of the patients, vice versa. Patients who were unable to read aloud at all (as confirmed beforehand using the CAT) performed the entire task (both lists) using inner speech alone.

For both populations, scoring of the inner speech task was based on the judgement given to a word pair, with possible answers being correct or incorrect. Hence, every pair judged incorrectly was scored as one error. Overt speech scores represent reading abilities. Participants were asked to read the words aloud and each word was scored separately, so every word read incorrectly was scored as one error. If the participant happened to read both words of a pair incorrectly, this would be scored as two errors. When a self-correction occurred, the word was scored as correctly read aloud.

Prior to the test, patients were given instructions and then a practice which included a minimum of ten items (more, when needed). In the practice, the experimenter read the items aloud first, and the patient had to give his/her judgement. This was done in order to confirm that the patient understood the task and had no significant receptive phonological impairment. As explained in chapter 2, healthy volunteers performed the entire task (both lists) using inner speech alone and immediately afterwards were asked to read aloud all the words. This allowed to record natural (mostly regional) variability in pronunciation of the words. Scoring of the tasks was identical to that described in chapter 2.

3. Results

Two patients (NET and WS) understood the word homophone task but not the rhyming task. The non-word homophone task proved to be especially difficult, with only 20 patients completing this task. Therefore, results of the word homophone task are presented for 27 patients, of the rhyming task for 25 patients, and of the non-word homophone task for 20 patients. For a list of tests completed by each patient see table 3.1. All patients completed the CAT and the cognitive test battery. For results of cognitive and language tests see table 3.3.

One of the healthy volunteers performed at chance level on some of the homophone judgement tests (homophone list 1: 42% correct, and non-word homophones list 1: 30% correct). Therefore, data from this outlier were excluded. Kolmogorov-Smirnov tests indicated that the data deviated significantly from the normal distribution for all tests (K-S test, $p<0.01$). Therefore, non parametric tests were subsequently used.

Table 3.3. Performance on the Comprehensive Aphasia Test (CAT) and the Apraxia Battery for Adults (ABA)

Test	Semantic memory	Semantic fluency	Phonological fluency	Recognition memory	Auditory comprehension	Reading comprehension	Repetition	Naming	Writing	Verbal Picture Description	Oral apraxia[1]
Range / Max[2]	10	unlimited	unlimited	10	66	62	60	58	76	unlimited	
Minimum[3]	8	13 (total fluency)		8	54	50	49	51	64	33	
AE	10	5	2	9	47	54	0	44[4]	65	NA	mild
AP	10	11	4	9	61	57	53	51	62	24	none
AS	10	9	4	10	66	58	56	56	74	34.5	mild
BA	10	21	22	10	66	62	60	58	76	110	none
BBS	7	1	0	8	53	53	8	16	58	NA	moderate
DB	10	17	8	9.5	61	61	58	53	76	36	none
DH	8	5	2	10	40	42	41	34	50	18.5	unknown
GH	10	4	2	10	61	54	57	48	60	49	severe
HD	9	0	0	10	51	44	31	15	25	16	none
IB	10	9	5	10	60	61	56	53	75	40.5	none
IC	10	3	2	5	62	42	14	9	42	NA	severe
JFI	10	4	4	6	61	56	30	43	68	18	mild
JFO	5	0	0	6	31	18	7	13	0	NA	severe
JHH	10	7	2	9	60	58	52	47	67	33	moderate
JS	9	9	5	10	61	61	60	54	76	57.5	none
LM	10	16	5	7	56	57	60	53	73	61	none
MM	9	10	4	10	63	60	60	54	76	38.5	none

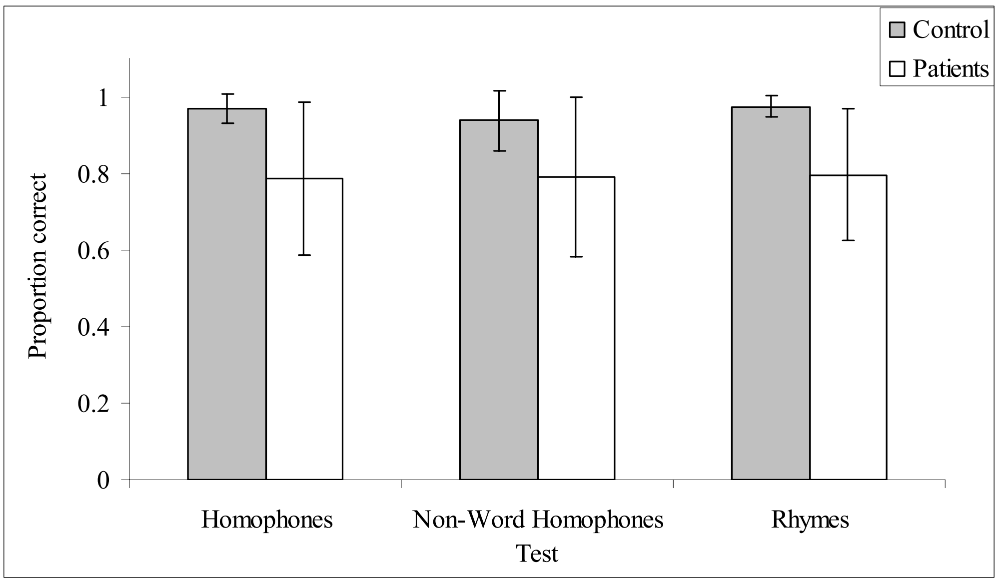
Test	**Semantic memory**	**Semantic fluency**	**Phonological fluency**	**Recognition memory**	**Auditory comprehension**	**Reading comprehension**	**Repetition**	**Naming**	**Writing**	**Verbal Picture Description**	**Oral apraxia1**
NET	10	7	1	10	56	51	17	10	48	13	moderate
PT	10	3	1	9	58	55	36	46	37	96	none
RB	10	1	1	10	42	44	49	15	43	1.5	severe
TH	10	7	7	9	60	58	48	52	75	56.5	moderate
WS	9	2	0	10	55	46	32	40	48	13	mild
CB	10	12	2	9	61	58	58	54	75	39.5	none
DR	10	8	4	10	66	59	60	54	76	41	none
EB	10	13	5	10	55	50	57	47	59	23	mild
CR	10	20	10	10	65	58	60	54	67	49	none
AD	10	8	4	10	62	58	26	24	74	15	moderate

[1] Oral apraxia was tested using subtest III of the ABA. Severity of apraxia was defined as: 0-30 severe; 30-40 moderate; 40-50 mild; 50 none.
[2] Possible ranges of scores or maximum scores are indicated for each sub-test.
[3] Minimum = minimum score of normal performance. Any lower score is considered impaired.
[4] Patient named by writing.

3.1 Do stroke patients with aphasia have impairments of inner speech?

The two groups (patients and controls) did not differ significantly in age or years of education (independent t-test, p>0.05 for both). Figure 3.1 shows the average score for each group in each of the three inner speech tasks.

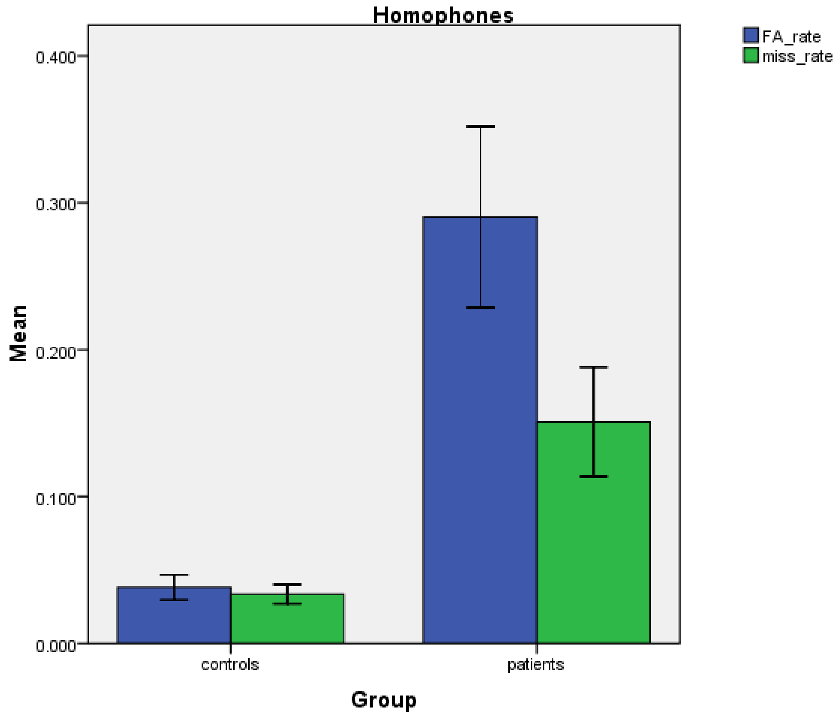
Figure 3.1. Average performance of the patient and control groups on the inner speech tasks. Error bars represent standard deviation.

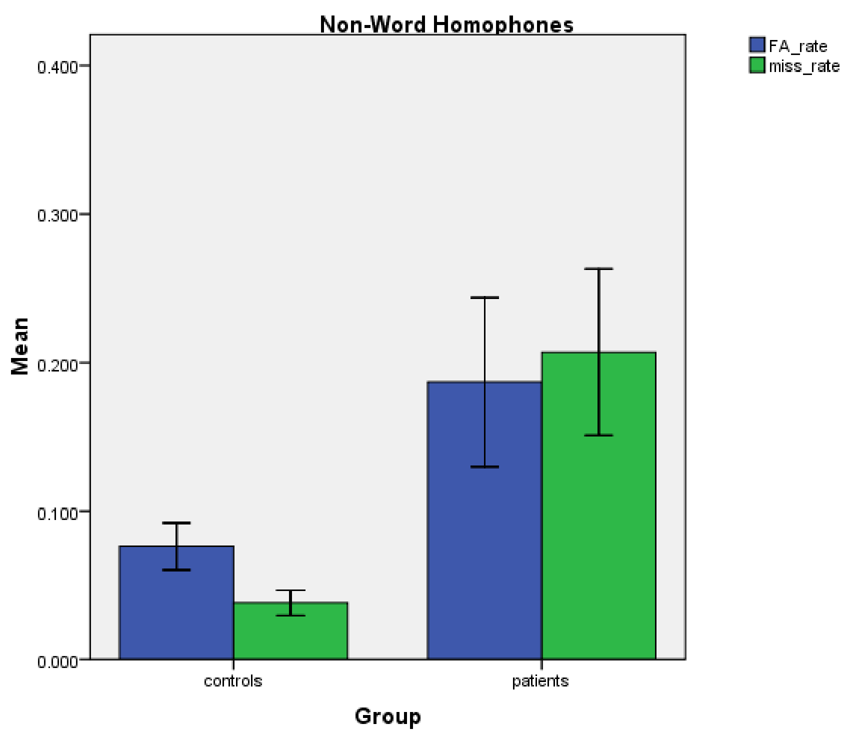


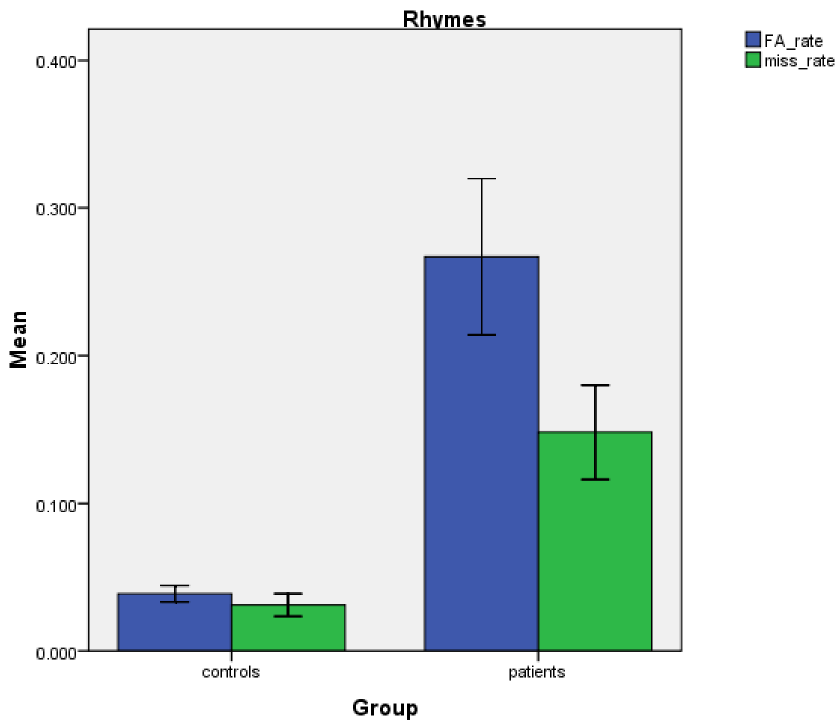
Overall performance for the two groups differed significantly in all three inner speech tasks (Mann-Whitney test, p<0.05 for all three tasks). In other words, the patients as a group were impaired relative to the controls in all three tasks. It is also clear from the graph, that the variance in performance in the patients group was higher than the variance in the control group.

It was examined whether any of the participant groups had a bias towards choosing one answer over the other ('yes' versus 'no'), and if such bias is found, whether it was different between the two groups and the various tasks. For this analysis, False Alarms (FA) and 'misses' were calculated. False Alarms are cases in which the participant answered 'yes' on a trial when the correct answer is 'no'. A 'miss' was recorded when a participant answered 'no' to a trial for which the correct answer is 'yes'. Therefore, 'FA rate' is the number of FAs out of the total number of trials in which the correct answer is 'no'. 'Miss rate' is the number of misses out of the total number of trials in which the correct answer is 'no'. Figure 3.2 shows the FA rate and the Miss rate, for each group and each condition.

Figure 3.2. FA rates and Miss rates for homophone judgement (top), non-word homophone judgement (middle), and rhyme judgement (bottom). Error bars represent standard error.







While controls had a bias towards answering ‘yes’ only in the non-word task (two-tailed paired t-test, t=2.555, p=0.013 for non-word homophone judgement, p>0.05 for the word homophone judgement and the rhyme judgement), patients showed a bias towards answering ‘yes’ only in the word task (two-tailed paired t-test, t=2.096, p=0.046 for the homophone judgement; t=2.307, p=0.030 for the rhyme judgement, p>0.05 for the non-word homophone judgement). There was no bias towards answering ‘no’ for any of the groups in any of the conditions.

If the inner speech tasks are used as a diagnostic tool, a cut-off score needs to be determined. A cut-off score of z=-2 was calculated, which equals 0.888, 0.778 and 0.918 correct for word homophones, non-word homophones, and word rhymes, respectively. Table 3.4 depicts the number of patients and controls who performed below the cut-off score in each test.

Table 3.4. Patients' performance on inner speech tasks relative to the cut off score of the different tests

Test	Cut off score (z=-2) Percent correct	Number of patients performing below the cut off score	Number of controls performing below the cut off score
Homophones words	88.8%	13 (48%)	1 (3.8%)
Homophones non-words	77.8%	8 (38%)	1 (3.8%)
Rhymes words	91.8%	15 (60%)	1 (3.8%)

For each patient and each test, it was calculated whether the score was significantly different from chance (50% correct), using the chi squared test with a p-value of 0.05. Table 3.5 presents the convergence of these two statistical tests. It can be seen that for many patients performance on the two tests converged: patients in group 1 performed significantly above chance and above the cut-off score, while patients in group 3 performed at chance level and below the cut-off point. However, there are some patients (group 2) whose tests do not converge: either scores are higher than chance but can still be considered impaired when using the cut off score, or vice versa.

Table 3.5. Number of patients in each of the groups, according to performance on the inner speech tasks (percentage out of total number appears in brackets).

Test	Group1: sig χ^2; z≤-2	Group 2: Mismatch between χ^2 and z score	Group 3: ns χ^2; z≥-2	Total number of patients
Homophones words	9 (33%)	4 (15%)	14 (52%)	27
Homophones non-words	8 (38%)	1 (5%)	12 (57%)	21
Rhymes words	3 (12%)	12 (48%)	10 (40%)	25

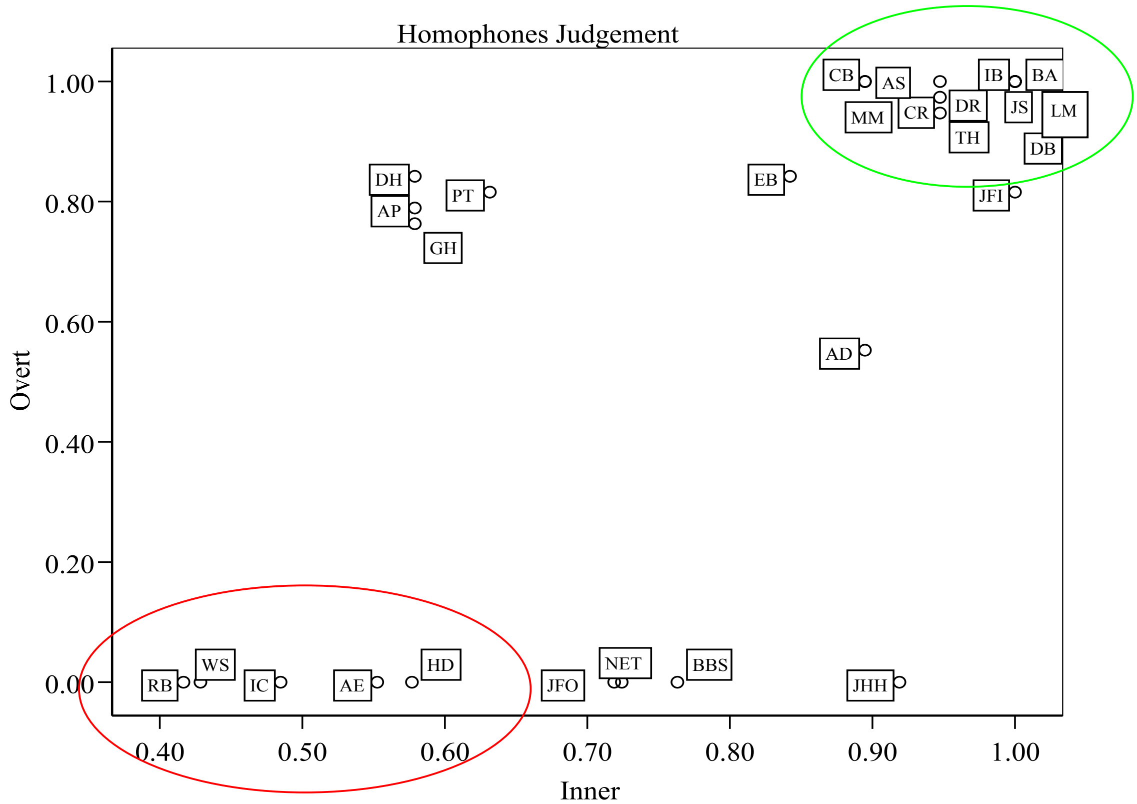
In summary, the patients as a group performed significantly less well than the controls, and demonstrated a wide range of impairments and abilities of inner speech.

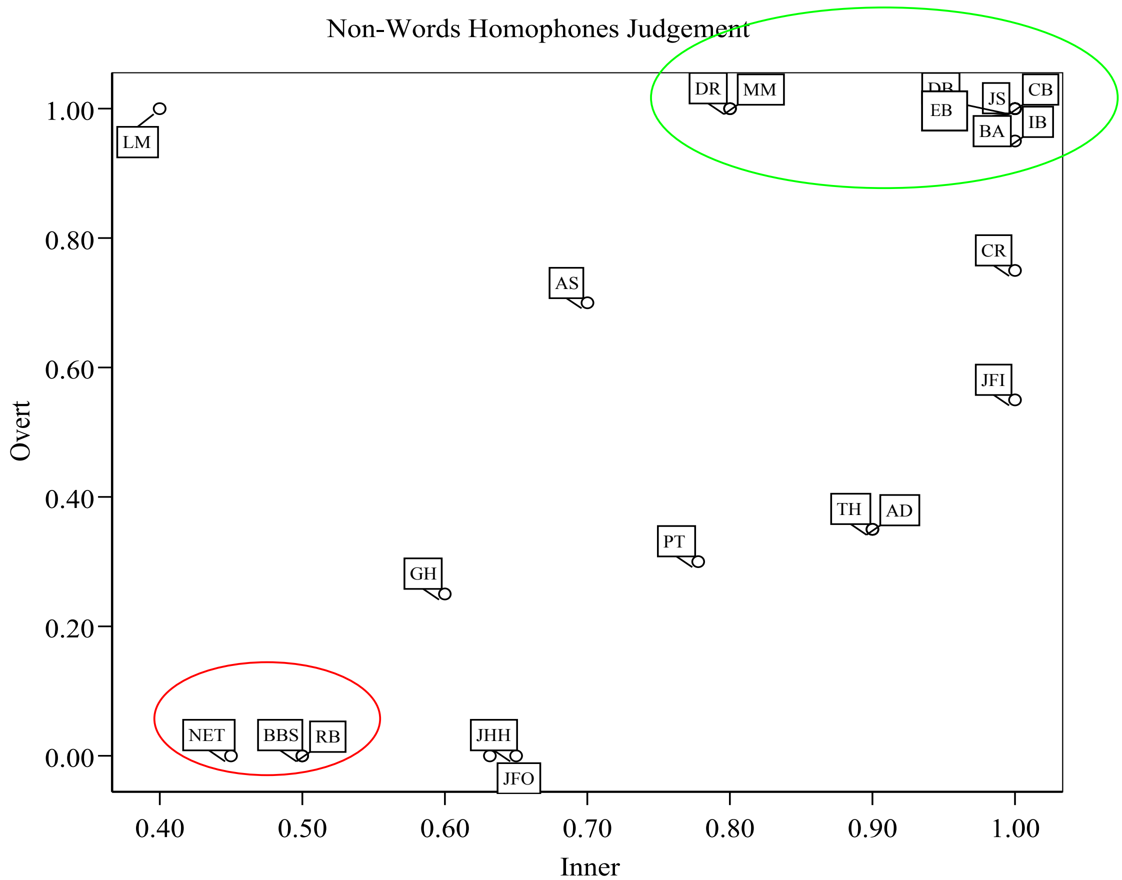
Lastly, it was examined whether inner speech performance could be solely explained by overt speech abilities and auditory comprehension. A regression analysis showed that 76% of the variability in performance on the rhyme judgement could be explained by the variability in speech production ($p<0.001$, scores for reading words aloud were taken from the CAT), and this model could not be significantly improved by adding speech comprehension (written word-picture matching task, scores taken from the CAT) as a predictor ($p=0.84$). Similar results were found when examining performance on the homophone judgement (54% of the variability was explained by word production performance, $p=0.015$, and the model could not be significantly improved by adding auditory word comprehension as a predictor, p=0.62). While overt speech abilities could significantly explain the variability in inner speech abilities, auditory comprehension did not significantly improve the model. This might be because auditory comprehension involves mainly semantic processing while inner speech and overt speech, mainly phonological processing. A better predictor related to comprehension might be the number of phonological errors made in the auditory comprehension task. However, since many patients did not make any phonological errors in comprehension, this variable could not be used in the regression analysis.

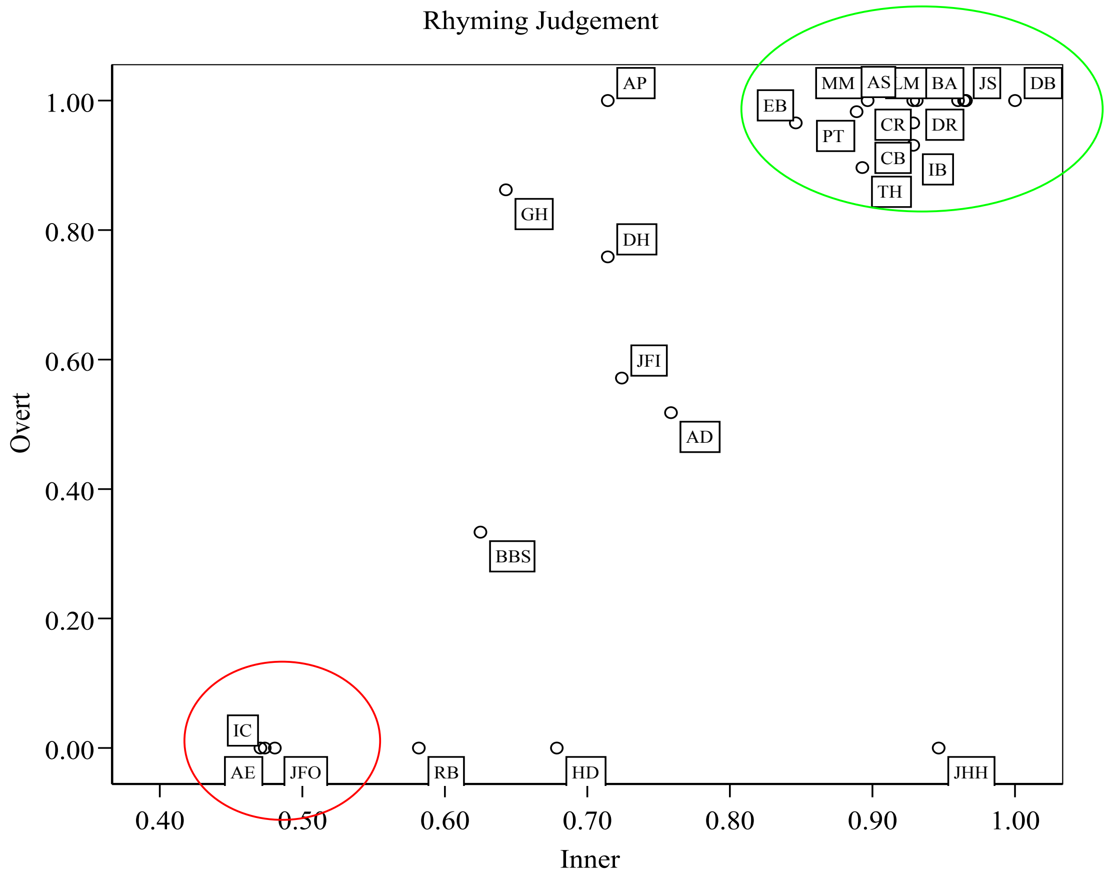
3.2 Do patients have dissociation between inner and overt speech abilities?

To investigate this question, correlations between patients' scores on overt and inner speech tasks were examined. Scores on overt tasks reflect the proportion of words read out correctly while scores on inner speech tasks reflect the proportion of pairs judged correctly using inner speech. Correlations between inner and overt speech were significant for all tests (one-tailed Kendall's Tau=0.60, $p<0.001$ for word homophones; Kendall's Tau=0.54, $p=0.001$ for non-word homophones; Kendall's Tau=0.64, $p<0.001$ for word rhymes). Figure 3.3 shows each patient's performance on the inner and overt speech tasks for each of the three tests. A cluster analysis revealed two main clusters in each task: patients with normal performance on both tasks (green circles) and those with severely impaired performance on both tasks (red circles). The high correlations mentioned above are therefore largely driven by these two extreme groups. All healthy volunteers (whose performance is not shown in these figures) fell into the green cluster in all tasks. While it is clear from figure 3.3 that many patients performed very poorly or very well on both tasks, some patients showed a discrepancy in performance between the two tasks. Furthermore, although some cases showing better inner speech than overt speech performance could be explained by the presence of motor speech deficit (apraxia), some patients in this subset do not have clear motor speech deficits.

Figure 3.3. Relation between inner and overt speech for each of the three tests: words homophone judgement (top), non-word homophone judgement (middle) and rhyme judgement (bottom).



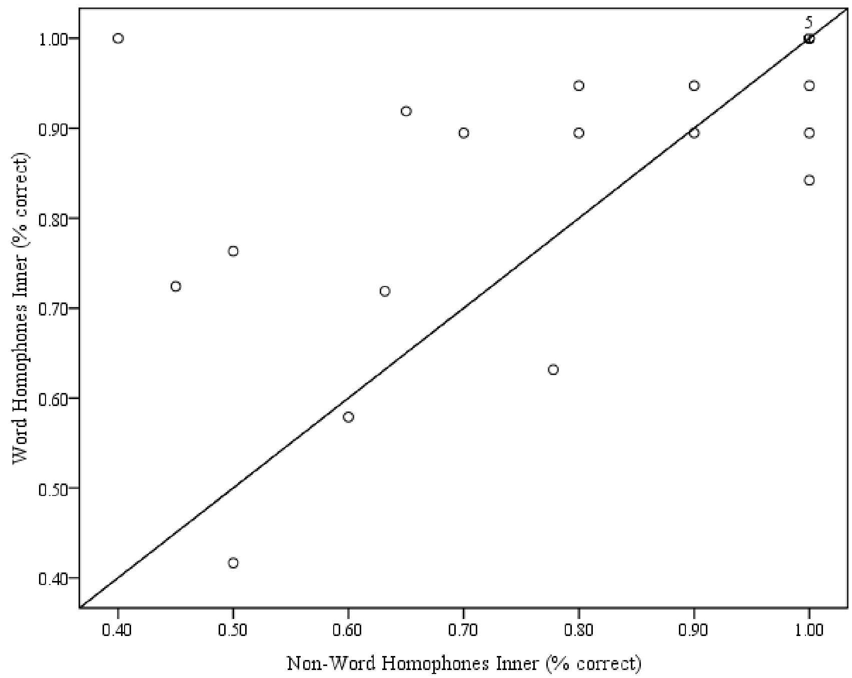

Non-Words Homophones Judgement
Overt
Inner
1.00
0.80
0.60
0.40
0.20
0.00
0.40
0.50
0.60
0.70
0.80
0.90
1.00
LM
DR
MM
JS
CB
EB
IB
BA
CR
AS
JFI
TH
AD
PT
GH
NET
BBS
RB
JHH
JFO

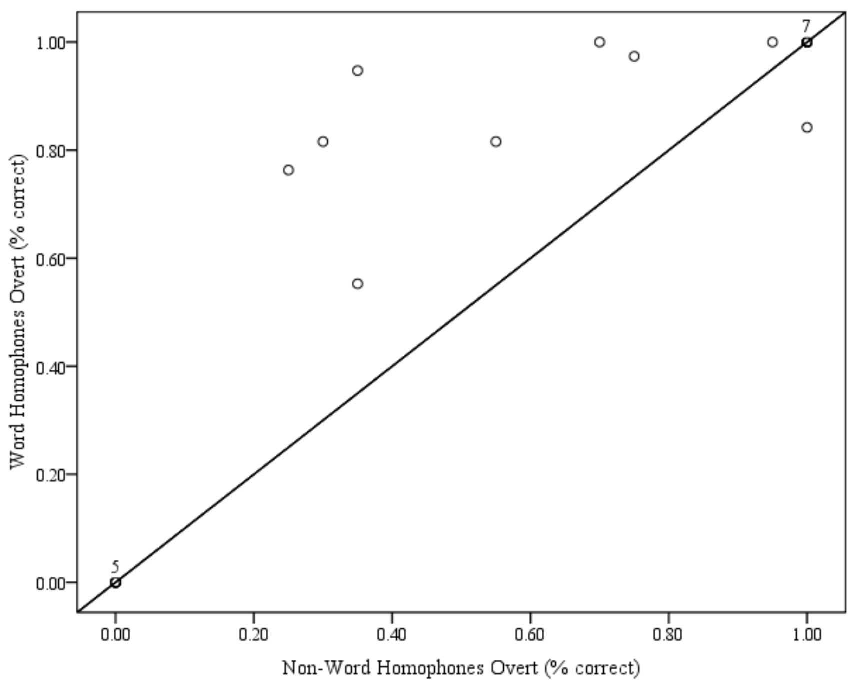

Rhyming Judgement
Overt
Inner
1.00
0.80
0.60
0.40
0.20
0.00
0.40
0.50
0.60
0.70
0.80
0.90
1.00
AP
MM
AS
LM
BA
JS
DB
EB
PT
CR
DR
CB
IB
TH
GH
DH
JFI
AD
BBS
IC
AE
JFO
RB
HD
JHH

3.3 Differences between Words and Non-Words

Next, performance on the word homophone task and the non-word homophone task was compared. Figure 3.4 shows the relation between the performances on the two tasks (n=21).

Figure 3.4. Performance on the word homophone task and the non-word homophone task using inner speech (top) and overt speech (bottom). The straight line represents equal performance on the two tasks. The number of cases represented by each dot is notated, when different from 1.





It is clear from the graphs that most patients fall on or above the line. That is, their performance on the word task is better or equal to their performance on the non-word task. For overt speech, scores on the task involving words were significantly higher than scores for the task involving non-words (one tailed Wilcoxon Signed Ranks Test, z=-2.43, p=0.008). For the inner speech condition, the difference only approached significance (one tailed Wilcoxon Signed Ranks Test, z=-1.45, p=0.074).

3.4 Case Studies

The behavioural profiles of specific patient were further explored. This included two patients who are representative of the impaired and well recovered groups, as well as other patients who displayed a substantial discrepancy between inner and overt speech. Performance on the language tests for all patients is presented in Table 3.3.

DB's performance on both the inner speech and overt speech tasks was at a normal level. Z scores for the tasks were all above the mean (inner speech tasks: z=0.816 word homophone judgement, z=0.798 non-word homophone judgement, z=0.962 rhyme judgement; overt speech tasks: z=0.283 homophonic word reading, z=0.537 homophonic non-word reading, z=0.354 rhyming-word reading). She reports that she can hear her own voice in her head when performing the inner speech tasks.

IC has speech production impairments which are partly a result of a severe speech apraxia. The inner speech tasks have the potential to reveal whether IC's failure to read aloud is due only to his speech apraxia or whether, in addition, he is unable to access and analyse the phonological properties of written material. His inner speech results clearly show that IC cannot access the sounds of written words: performance was at chance level in both of the inner speech tasks that he attempted (word homophone judgement: z=-12.174, chi squared: p=0.136; rhyme judgement: z=-18.247, chi squared: p=0.102). Hence, IC's inability to read aloud is not only a result of speech apraxia but also a result of severe language impairment.

In the rhyme judgement task three patients, DH, AP and GH, showed better preserved overt speech compared to inner speech. Can this deficit in rhyme judgement be attributed to impairment in working memory?

DH has no obvious signs of apraxia. On the rhyming task DH is somewhat impaired, with a z score below -2 (z=-9.404). However, his performance is above chance (71.4% correct, chi squared: p=0.01). Sentence repetition is impaired as well (score 6/12). When presented with a sentence, DH can only reproduce the last part of the sentence, and sometimes produced phonological errors. For example, from the sentence 'The man went and shut the window', DH could only repeat the word 'window'. Trying to repeat the sentence 'The children listened as the teacher read the story', DH replied: 'read the sory'. Auditory sentence comprehension is impaired as well (score 16/32), and performance is slightly lower than that achieved for written sentence comprehension (score 20/32). Therefore, impairment in auditory comprehension might be partially attributed to working memory impairments as well. Lastly, it should be noted that DH also performed poorly in the homophone judgement task (57.9% correct, z=-9.801, chi squared, p=0.14). Hence, DH's behavioural profile suggests that he is suffering from impairments in both working memory and inner speech, and his performance on the rhyming task is therefore likely to be affected by both impairments.

AP's performance on the inner speech tasks is impaired (homophone judgement: 57.9% correct, chi squared: p=0.14, z=-9.801; rhyme judgement: 71.4% correct, chi squared: p=0.01, z=-9.404; non-word homophone judgement was not performed). He shows no difficulty with sentence repetition (score 12/12), and auditory and written sentence comprehension are equal and only slightly impaired (score 28/32 for both). From this data, it seems that AP has no severe deficit of working memory, but he has a clear deficit of inner speech.

GH performed close to chance level on the rhyming task (64.3% correct, chi squared: p=0.048, z=-11.995). However, sentence repetition is normal (score 12/12). Auditory sentence comprehension is just below normal (score 27/32), but is slightly better than or equal to written sentence comprehension (score 24/32). These results do not definitively establish an impairment of working memory. On the homophone judgement GH performed at chance level (57.9% correct, chi squared: p=0.14, z=-9.801). In summary, GH, like AP, seems to have no severe deficit of working memory, but has a clear deficit of inner speech.

Lastly, in the non-word homophone judgement, **LM** stands out as having intact overt speech but highly impaired inner speech. Rhyme and homophone judgements of words were normal, both for inner speech and overt speech (inner speech: homophone judgement, success rate=100%, chi squared: p<0.001, z=0.815; rhyme judgement, success rate=96%, chi squared: p<0.001, z=-0.048. Overt speech: homophones; success rate=100%, z=0.283; rhymes; success rate=100%, z=0.354).

Nevertheless, while performing the inner speech tasks with words successfully, she was completely unable to do so with non-words. Remarkably, she could read the non-words aloud (overt speech task of non-word homophones, success rate=100%, z=0.537). She reports that she can clearly pronounce words in her head but cannot pronounce non-words in her head at all. Her normal performance on sentence repetition (12/12) suggests that she does not have working memory impairment. Moreover, performance on auditory sentence comprehension, although slightly impaired, is equal to written sentence comprehension (26/32 and 27/32 respectively). Again, this suggests that impairments in sentence comprehension cannot be attributed to working memory deficits.

3.5 Do patients rely on inner speech or overt speech when the two differ?

In order to address this question, the relation between patients' overt word reading and homophone/rhyme judgements was examined on an item by item basis. Note that in the previous analysis the scores for overt performance were from reading aloud. In the current analysis, the scores for overt performance are those of the judgement given when reading the word pairs aloud.

For each patient, all items were classified according to both reading (correct or incorrect) and judgement (correct or incorrect). These data were available for 16 patients, as 8 patients performed the task using inner speech only and for 3 more patients exact pronunciation during reading aloud was not recorded.

When a homophone or rhyme judgement is made following correct reading aloud of the words, it cannot be determined whether the judgement was based on inner or overt speech, regardless of the judgement given. This is because, in reading aloud, both overt and inner speech processes are available to the reader. When a word is read incorrectly, on the other hand, there are two possible results: the first is when a word is read incorrectly, but a judgement made based on inner speech should be the same as one made based on overt speech (inner = overt). For example, the 'break-creak' pair in the rhyme judgement task does not rhyme. If read as 'rake-reak', it still does not rhyme. Hence, the error in reading does not change the judgement which should be given: judgements based on inner or overt speech are the same. The second case is when words are read incorrectly, but a judgement made based on inner speech should be the opposite to the one made based on overt speech (inner ≠ overt). For example, the pair 'wart-fort', which appeared in the rhyme judgement task, is a rhyming pair. When read erroneously as ‘woot-sot’, the pair becomes a non-rhyme. Hence, judgement given based on inner and overt speech should differ. This case is the most interesting to the experimental question, since only for such items it is possible to determine whether the judgement given by the patient is based on inner speech or not. Unfortunately, although this analysis is relatively simple and potentially revealing, the data set did not generate enough items of the appropriate kind in order to reach clear conclusions about individual patients. The patients who had the largest number of relevant items suggest some preferences. JFI and AD made many judgements based on inner speech (5 and 9 respectively). Interestingly, both patients showed clear awareness of overt errors together with an effort to correct them. Their inability to correct overt errors in most cases is largely due to speech apraxia. Hence, for these two patients inner speech is better preserved than overt speech and therefore they correctly rely on the former rather than the latter for speech monitoring purposes.

GH performed at chance level on all of the inner speech tasks. When reading the words aloud, she still performed at chance level. In many items on which an error was made (either by giving an incorrect judgement or reading aloud incorrectly), the judgement given did not match the way the words were pronounced out loud (22 out of 58 items in all tests together). For example, pairs which were read correctly were still judged wrongly. It is therefore suggested that although her inner speech is impaired, she continues to rely on it for error monitoring. As a result, when producing accurate overt speech, she fails to use this information correctly.

PT made more errors when the task was performed using inner speech alone, compared to the condition when overt speech was added. Moreover, PT did not show awareness of his errors in spontaneous speech or during task performance, and therefore did not try to correct them. Taken together, these data suggest that PT relies more on overt speech, even when it is erroneous.

4. Discussion

This study investigated inner speech in post-stroke aphasia and its relation to overt speech production. The results suggest that while most patients have similar overt and inner speech abilities, a proportion demonstrate very discrepant levels of performance on these two aspects of language ability. The results have implications for understanding normal language processing as well as for understanding impairments in aphasia.

4.1 Inner Speech Abilities in Aphasia

Looking at inner speech alone, the patients as a group were clearly impaired relative to the control group. However, the variability within the patient group was large, with some patients performing at a normal level while others demonstrated clear impairment. Two different ways of analysing the results of the inner speech task were explored. The first method tested whether each patient's score was significantly different from chance. The second method is similar to the one used in standard neuropsychological tests, i.e. the use of a cut-off score (Berg 1948; Wechsler 1955; Folstein, Folstein et al. 1975; Burgess and Shallice 1997). In this study, the cut off score was calculated as 2 standard deviations below the mean of the normal, age- and education-matched, control group. The two methods gave slightly different results. Although the majority of cases were classified as either falling within or outside the normal range by both methods, a sub-group scored around 70% correct – which is significantly above chance but below the cut-off score. This was evident especially in the rhyming task, with 48% of patients showing this 'mismatch' (see table 3.5). This observation emphasises the problem with describing inner speech as either impaired or preserved. It is clear that, when testing patients' inner speech, one should treat the data as continuous rather than discrete. Previous studies looking at inner speech failed to describe this variability, for various reasons. First, in some studies, patients clearly fell in one of the two extreme groups (intact inner speech as in Nebes (1975); complete absence of inner speech as in Levine et al. (1982); or with representatives of both extreme groups as in Feinberg et al. (1986)). Secondly, some authors used only a cut-off score, which concealed the variability. For example, Pratap (1987) defined electroencephalography (EEG) signals as either normal or impaired. Lastly, single case studies (Baddeley and Wilson 1985) cannot explore the range of abilities within a patient population.

A response bias was observed in some of the tasks. While the patients showed a bias towards giving a positive answer in the word tasks, the controls showed a small bias towards giving a positive answer only in the task using non-words. There was no bias towards giving a negative answer in any of the tasks, by any of the populations.

A response bias is often a result of noise in the stimuli. Healthy volunteers might have found the non-word task more 'noisy'. This is because the pronunciation of non-words is supported only by the phonological system, while pronunciation of words is supported by both the phonological and the semantic system. Consequently, a non-word task has higher ambiguity than a word task. It is well known that, in general, people perform less well on tasks using non-words than on tasks which use words. Since the volunteers' performance on the non-word task was equal to that on the word tasks, the ambiguity in the task might have come into play by producing this small response bias. The patients' bias potentially has a different source. Patients with language production impairments find both tasks ambiguous to some extent. Moreover, many patients are unsuccessful in correcting or even detecting their own errors, but still they know that they produce errors. This might create an inherent bias in their response, in which the patient knows that he often gets words wrong, and therefore attributes the non-rhyming or non-homophony to his own language impairment, and attempts to 'correct' it by answering 'yes'. However, one would expect this bias to occur in the non-word task as well and it is not clear why the patients showed this bias only in the word task.

4.2 Relation between Inner and Overt Speech

Figure 3.3 shows that patients' performance was distributed across the entire range of scores. Nevertheless, for all tasks, half or more of the patients fell into one of two clusters: either high (green circles) or low (red circles) performance on both reading aloud and silent judgement tasks. Patients with poor performance on both tasks might have deficits in accessing the phonological form of the word from writing, as is the case in deep alexia (Benson and Ardila 1996). Impairment at this level will be evident not only when the patient reads aloud, but also during silent reading when the task requires phonological access. Some idea as to whether patients' inner speech impairment can be explained by a reading impairment comes from the scores on the reading comprehension tasks of the CAT. After examining the reading performance of patients who performed poorly on the inner and overt speech tasks, it was found that these patients differ in their behavioural profile. Some patients showed a clear impairment in reading comprehension, which cannot be attributed to a comprehension disability per se, since their performance on the auditory comprehension task was normal (for example, patients IC and WS) or substantially better than performance on the reading comprehension task (for example, patient JFO). For these patients impairment in the inner speech tasks can be at least somewhat attributed to their reading impairment, and therefore they might perform better with inner speech tasks which use pictures or spontaneous speech. Still other patients have no reading comprehension disability (for example, patients NET, JHH, BBS and AE). For these patients it is less likely that their failure in the inner speech tasks can be fully attributed to the fact that written material was used.

For the patients who performed the inner speech tasks but did so unsuccessfully it was sometimes difficult to determine what guided their answers. The data show that some of these patients based their answers on the orthographic similarity between the two words. These patients judged ‘town’ and ‘gown’, but also ‘give’ and ‘five’, as rhyming pairs. Similarly, they judged ‘pea’ and ‘play’, but also ‘chair’ and ‘bear’, as non-rhyming pairs. Others might have used implicit knowledge about the correspondence between the orthographic and phonological form of words, to make their judgement. However, this strategy was not beneficial since many words were irregular. Altogether, it seems that the group of patients who performed poorly on both the inner and the overt speech tasks is heterogeneous with regard to the strategies they used to solve the task and their underlying disabilities.

Of greater interest, of course, are the patients with discrepant levels of success in the two tasks. In the word homophone judgement task, patients AP, DH, GH and PT performed at chance level, or close to it, on inner speech judgements, but showed relatively good ability to read words aloud. AP, DH and GH also showed the same pattern in the rhyming task. LM was characterised by this pattern with regard to non-word homophone judgement only. This impairment cannot be attributed to a deficit in phonemic discrimination since all patients were successful in rhyme and homophone judgements when pairs were read aloud by the examiner. In some cases, such as DH, a discrepancy between inner and overt speech may be attributable in part to working memory deficits: judging the rhyme status of two words uttered by someone else requires some capacity of working memory; but the burden on working memory is far greater if one must first derive the phonological representations for oneself before comparing them to judge their relationship. And indeed, many aphasic patients have difficulties with working memory (Caspari, Parkinson et al. 1998; Wright, Downey et al. 2007; Baldo, Klostermann et al. 2008), and this deficit can, in some cases, explain certain comprehension impairments presented by them (Miyake, Carpenter et al. 1994; Caspari, Parkinson et al. 1998).

Although this study did not include a thorough examination of working memory, the assessments of sentence comprehension and repetition do offer at least some indication of this ability; and by these measures, some of the patients with poorer inner- than overt-speech ability (such as AP, GH and LM) do not seem to have much of an impairment in this domain. Moreover, homophone judgement should not be influenced by working memory (Howard and Franklin 1990). A number of ways are suggested to explain how inner speech impairment can be accounted for by current models of speech production. Firstly, as discussed in the pervious chapter difficulty with inner speech tasks relative to overt speech can be attributed to lack of motor and overt auditory feedbacks during inner speech. And indeed, as mentioned in chapter 2, healthy participants vocalise more as the difficulty of reading material increases, and have lower comprehension scores when prevented from vocalising or even subvocalising (Hardyck and Petrinov 1970). Oppenheim and Dell's (2008) study, suggesting that inner speech is phonetically impoverished in comparison to overt speech, also supports the finding that inner speech performance is worse than overt speech performance. In summary, patients who show impairment in inner speech relatively to overt speech replicate the results from healthy controls and explanations to this phenomenon can account for both populations.

Levelt (1999) and others[11] (Monsell 1987; Martin, Lesch et al. 1999) suggest that inner speech is produced by using the production as well as the comprehension system. Accordingly, inner speech deficit can arise from impairment to the production system, comprehension system or the connection between them. If the deficit is in the production system then overt speech should be impaired as well. This is the case for the many patients who show impairments on both inner and overt speech tests (figure 3.3, red circles). A deficit in the comprehension system will impair both comprehension and inner speech. However, in this patient cohort most patients have impairment in speech production and only very few have comprehension deficits. Therefore no patients were found who are impaired only on comprehension and inner speech. However, such cases can be probably found. Lastly, inner speech impairments can arise from difficulties in the transfer of information from the output phonological

[11] A comprehensive discussion of these models can be found in chapter 1.

store to the input phonological store. Such impairment would not necessarily cause a deficit in overt speech, since overt speech production does not require activation of this pathway. It would also not impair spoken or written comprehension. AP and GH have relatively preserved comprehension as can be seen in both the reading and the spoken tasks. Lastly, impairment in transferring information between the output and the input phonological stores will not necessarily cause a deficit in working memory that can be observed in the tasks used in this study. Both tasks (the sentence repetition and the sentence comprehension) had a strong semantic component. Semantics can support working memory (Martin et al., 1999), therefore allowing patients to perform these tasks successfully. The rhyme and homophone judgement had a very minimal semantic component and as a result, patients who have impairment in inner speech could not perform the tasks as successfully. The contribution of semantic support to inner speech processing is especially evident in the case of LM, who is further discussed below.

In summary, models which suggest a separation between the output and the input phonological stores are compatible with the observation that some patients have preserved overt speech together with impaired inner speech, over and above working memory deficits.

Looking at connectionist models, such as Dell's (1986), the findings are harder to account for. Feedback connections, which might be responsible for inner speech during error monitoring, have a role in increasing efficiency and accuracy of output (Dell 1986; Postma 2000; Vigliocco and Hartsuiker 2002). If feedback is compromised, causing the observed deficit in inner speech, overt speech could be affected as well. However, this is not the case for some of the patients. Therefore, connectionist models in the simple form presented above have some difficulty in explaining the patterns of deficits and abilities observed in this study.

One of the more interesting cases in the data is patient LM, who has complete absence of inner speech when non-words are involved, but perfectly intact inner speech when the task uses words. A model which fully dissociates the comprehension and production systems at the phonological level cannot easily account for this finding. If the connection between output and input phonology is impaired, it would be expected that inner speech will be affected for both words and non-words. The discrepancy suggests some dissociation between a lexical and a sub-lexical route. Incorporating feedback connections into the model can explain the data. Accordingly, LM's connection between output and input phonology is severely disrupted, which does not allow her to create inner speech using this direct route. However, feedback connections from the phonological output store, into the semantic system, and between the semantic system and the phonological input store, allow LM to produce inner speech when words are involved, by using this indirect lexical route.

The other interesting cases are those who have a clear impairment in overt speech, but preserved or relatively preserved inner speech. In the current study, five patients (AD, JHH, BBS, NET and JFO) performed above chance in the homophone judgement, but had clear difficulties with reading the words aloud. AD, JHH and BBS showed the same pattern of results in the rhyme judgement as well, together with HD and JFI. Lastly, the homophone judgements of non-words uncovered three patients (AD, JFI, and TH) who had difficulty with reading out loud but above-chance performance on the task using inner speech. One obvious explanation for such discrepancy between inner and overt speech would be the presence of a motor speech disorder. Indeed, most of these patients have moderate to severe oral apraxia. In such cases the use of inner speech tasks could shed light on the linguistic impairments, revealing whether such impairments exist over and above the motor deficit. However, these tests are not routinely done with patients with aphasia. Regarding the patient HD who has no motor deficits but still shows the above dissociation, it is suggested that he has an impairment in translating the phonological code into the articulatory code required for the production of overt speech, a process described by many (Levelt 1989; Martin 2003; Guenther 2005; Hickok and Poeppel 2007). This deficit can also be described as an inability to maintain stable levels of activation (Wilshire 2008), resulting in abnormally rapid decay of information - as described for patient NC, in Martin et al. (1994). Anecdotal support for this account comes from JHH, who describes why he cannot pronounce aloud words that he can judge for rhyming or similarity. He claims that he can retrieve and hold the word in his head only long enough to perform the task, but when trying to say the word aloud the trace disappears: 'I have it… it just goes away, it is gone', a simple description of abnormally rapid decay of information.

If inner speech was simply overt speech without articulation, only one type of dissociation would be expected: impaired overt speech with preserved inner speech. The fact that the reverse pattern occurs as well suggests that such a description of inner speech cannot be accepted. Both types of models - those which emphasise the importance of the comprehension system in inner speech processing, as well as connectionist models which incorporate feedback within the production system - are challenged to account for the variability of behaviour in this patient population. It is proposed that the data in this study can be explained by describing inner speech as primarily a connection between output and input phonology, i.e. between the production and comprehension systems. However, when this connection is damaged, an indirect route, using feedback connections, can support lexical inner speech.

Language models usually discuss inner speech in relation to error monitoring. However, inner speech also has a role in reading, writing, thinking and other cognitive functions, and these processes are also addressed in experimental work. Vigliocco and Hartsuiker (2002) point out that inner speech, in the presence of simultaneous overt speech and in the absence of overt speech, might differ. Following their view, it is proposed here that future studies should differentiate between the two types of inner speech: conscious inner speech, which was tested in this study, versus unconscious inner speech, which might be the one responsible for on-line error monitoring. It is still debatable whether the latter is solely production-based or dependent on the comprehension system. However, differentiating the two types of inner speech might shed light on the nature of some of the disagreement in the literature today.

4.3 Differences between Words and Non-Words

A comparison of the pattern of results for the two tasks - one using words and the other using non-words - reveals a few clear differences. Firstly, scores on the task using non-words were more evenly distributed than scores for the word judgements. Secondly, and as hypothesised, most patients performed equally well, or worse, on homophone judgements for non-words relative to words. This finding was true in both the overt speech and the inner speech conditions. Moreover, some patients who could perform the tasks using words were unable even to attempt the non-word task. Furthermore, it is interesting to note that individual patients might show partial impairment for non-words only. Patients LM and TH performed within the normal range in both rhyme and word homophone judgements (for both inner and overt speech tasks) but, when it came to non-words, the two patients showed opposite patterns: while patient TH has an impairment of overt speech, patient LM had an impairment of inner speech only. These finding replicate the results by Wilshire (2008), who reviewed cases of patients who produce phonological errors in all types of tasks (naming, repetition and spontaneous speech). Her review shows that while errors are produced for both real- and non-words, patients produce less errors with real words than with non-words (Wilshire 2008). As discussed in chapter 2, all current models of speech production (see for example Paap and Noel 1991; Coltheart, Curtis et al. 1993; Martin, Dell et al. 1994; Martin, Saffran et al. 1996; Plaut, McClelland et al. 1996; Dell, Schwartz et al. 1997), despite their real differences, incorporate one or more sources of lexical support for phonology. Moreover, the connections between the phonological units of a word are well practiced and therefore have strong activation, stronger than that of non-words (Acheson and MacDonald 2009). Since words have these additional sources of activation, the chance of choosing the correct phonological unit for words increases in comparison with non-words, and accordingly, the chance of producing an error, decreases (Wilshire 2008). In cases with an abnormally fast decay of activation, as described above, activation of words will be renewed often and to a high extent, since it is supported by both semantic and phonological input and feedback, while activation for non-words will be renewed less often, or to a lesser extent, and is therefore more vulnerable to complete decay (Martin, Dell et al. 1994; Martin, Saffran et al. 1996). The inevitable consequence is the results seen in this and previous studies: unless performance on a task involving non-words is at ceiling, scores on tasks involving word exceeded those on non-words.

The finding that reading of non-words is in general more impaired than reading of real words is further supported by observations and classifications of aphasia. Patient presenting with phonological dyslexia have difficulty reading non-words despite relatively preserved ability to read real words (Shallice and Warington 1980). Patients with surface dyslexia on the other hand, have difficulty with reading irregular words despite relatively preserved ability to read non-words and regular words (Patterson and Morton 1985). Lastly, in deep dyslexia, reading ability is highly impaired for all materials (Marshall and Newcombe 1973). From this description it is clear that a case of preserved reading of non-words together with impairment of reading all real words does not exist in aphasia, to the best of our knowledge. This superiority of real words is further supported by the findings presented here.

However, as can be seen in figure 3.4, for a small number of patients performance on the non-word homophone task is relatively better than performance on the word homophone task. These differences might be a result of patients' sensitivity to linguistic variables such as word length or frequency of syllables. The two tasks, the one involving words and the one involving non-words, were not matched on these variables and in some cases these linguistic variables might have been more influential than the simple lexicality effect.

4.4 Clinical Implications

It is sometimes assumed that, when patients make overt speech errors, this reflects a deficit in all levels of phonological processing. These results suggest that this is often, but not always, true: in some cases, inner speech can remain relatively intact side-by-side with a marked deficit in overt speech. It is therefore suggested that tasks using inner speech might be used to improve diagnosis, especially in cases where patients have marked motor deficits. Even for patients without motor deficits, inner speech tasks can reveal speech capabilities that remain hidden when overt speech alone is examined. The opposite pattern, impaired inner speech together with preserved overt speech, has been found in this study as well. In these cases assessing overt speech alone will not give the clinician a full picture of the patient's deficits. It is therefore recommended to assess inner speech even when overt speech is preserved. Patients with such impairments might have other related difficulties which require further investigation (for example, verbal working memory or reading). Lastly, an analysis of reading and judgement errors, as was done in this study, can potentially reveal which strategy patients use for monitoring their speech.

In the introduction, a discussion of current diagnosis methods and their caveats was laid out. The work presented in this chapter aimed to address some of the problems with current diagnostic tests. Firstly, by using the CAT, some of the caveats of other tests are avoided. Specifically, the CAT is based on neurocognitive models of language processing, and it is reliable and valid (Swinburn, Porter et al. 2004). By using the CAT and the inner speech tasks, there was no attempt to classify patients into syndromes or otherwise. Rather, it is argued that when making diagnosis one should take into consideration the variability within the patient population. Classifying patients into syndromes, or into 'impaired' versus 'non-impaired' groups, stands the risk of overlooking valuable information regarding patients' underlying impairments. It was also argued in the introduction that most aphasia batteries overlook some important aspects of language processing (Simmons-Mackie, Threats et al. 2005). While it is acknowledged that this study did not address all aspects of language function, it presents an attempt to bring into light an important aspect of language, i.e. inner speech, which was largely overlooked in the past and is not routinely assessed by clinicians. Lastly, some authors argue that trying to infer about lesion from language data is ineffective, especially in light of current advances in brain imaging (Bruce and Edmundson 2010). In the next chapters, imaging data will be incorporated in order to build a more comprehensive picture of inner speech and its disturbances.

4.5 Conclusions and Future Directions

There were several limitations to this study. Firstly, the number of items in the inner speech tasks did not allow to determine whether patients tend to rely more on inner or overt speech for error monitoring. Moreover, all inner and overt tasks were based on reading. For patients with severe reading disabilities, it would be advantageous to develop tasks which do not require reading; for example, the use of pictures or automatic speech. Future research should also focus on properties of inner speech which were not explored in this study. This study compared words and non-words, and their influence on inner and overt speech abilities. It remains to be established whether other linguistic variables such as word frequency, phonological neighbourhoods or word length, influence inner speech abilities as well. Moreover, performance on the homophone and rhyme judgements differed for some, but not all, patients. Some studies suggest that these tasks have different requirements and therefore might be dependent on different mechanisms (Howard and Franklin 1990). In this study, the words used for the two tasks were not matched for linguistic variables, such as word frequency or number of syllables. Therefore, any comparison between the two tasks could have been essentially contaminated by these variables. Lastly, this study is a population study. It was aimed at identifying different profiles of behaviour with regard to inner speech, within the population of aphasic patients. Further investigation of some of the patients might shed light on their specific impairments and the relations between conscious inner speech and pre-articulatory inner speech.

It is concluded that patients with aphasia seem to be distributed across the entire spectrum of abilities of both inner and overt speech. For most patients, performance on one of these two aspects of language is a reliable predictor of the other; but the findings indicate that some patients can have relatively preserved inner speech with a marked deficit in overt speech and vice versa. Future studies can explore whether inner speech abilities can predict recovery of other language abilities, such as auditory or reading comprehension of texts. These results have important implications for the understanding of normal inner speech as well as inner speech in post-stroke aphasia and its relation to diagnosis and therapy. Lastly, by uncovering the neural correlates of inner speech, by combining lesion studies with functional brain imaging, future research may deepen our understanding of inner speech in both aphasic and healthy participants. The next chapter is therefore aimed at exploring some of the neural mechanisms underlying inner speech.

Chapter 4
A Voxel-Based-Lesion-Symptom Mapping Study of Inner Speech

A science of the mind must reduce [complexities of behaviour] to their elements. A science of the brain must point out the functions of its elements. A science of the relations of mind and brain must show how the elementary ingredients of the former correspond to the elementary functions of the latter."
William James, The Principles of Psychology (1890)

1. Introduction

1.1 Lesion Analysis Techniques

The anatomical correlates of the language system have been discussed in the western literature for more than 140 years, starting with the famous works by Broca and Wernicke. Thousands of studies discuss the anatomy of language, using numerous techniques, tasks, languages and populations. Until recently, two main methodologies dominated the scene of mapping symptoms to lesions. In the 'lesion-defined' technique, two groups of patients with different types of lesions (for example, perisylvian versus extra-sylvian, as in Rapcsak, Beeson et al. 2009) are compared on behavioural measurements. If one group is significantly more impaired on a specific behavioural measurement, then this behaviour can be mapped onto the relevant brain area. For example, Kuljic-Obradovic (2003) defined the different symptoms presented by aphasic patients with different subcortical strokes: thalamic, striato-capsular and white-matter lesions. They found that striato-capsular and white matter lesions were associated with a fluency deficit, while thalamic lesions were associated with impaired comprehension and naming (Kuljic-Obradovic 2003). Although widely used, this method has a few major caveats. Firstly, by looking at Regions of Interest (ROIs), one might overlook the importance of areas outside these regions, as well as the distinct importance of subareas within the region. Secondly, this technique requires two groups of relatively homogenous populations with regard to the site of stroke. The second technique groups patients according to behavioural deficit rather than lesion site. Using this method, one would compare two groups of patients with different behavioural abilities, examining whether they also differ in the site of lesion.

A study using this method managed to define areas involved in speech articulation by comparing patient with and without apraxia of speech (Dronkers 1996). This approach has some caveats, with the main one being that behavioural symptoms need to be defined as either preserved or impaired. Clearly, most behavioural measurements are more complex and therefore by examining data in a binary fashion, crucial and interesting data can be overlooked. Moreover, the cut off point can be arbitrary at times, resulting in different outcomes when using different cut off points. As a consequence, researchers often tend to look at groups of patients showing extreme behavioural profiles; patients with substantial impairment versus those with no impairment at all. By doing so, again, essential data are lost and the study cannot be easily generalised to the patient population not tested. With regard to inner speech, the disadvantages of treating inner speech as either preserved or impaired have been demonstrated and discusses in the previous chapter.

Building on this clear need for a new technique for lesion analysis, Bates et al. (2003) have developed the Voxel based Lesion Symptom Mapping (VLSM), which is an improvement on the traditional lesion symptoms mapping techniques described above. In VLSM, patients are divided into two groups according to whether they do or do not have a lesion affecting a specific voxel. Behavioural scores are then compared for these two groups, yielding a t-statistic for that voxel. The procedure is then repeated again and again for each voxel in the brain. This method deals with the two main caveats of previous techniques; the disadvantages of working with ROIs and the loss of information when using binary behavioural data.

A different method employs perfusion data for the mapping of behavioural deficits in acute stroke. This aims to overcome some of the main caveats of the traditional lesion-symptom mapping techniques. Firstly, approaches that examine only chronic patients have an inherent selection bias since usually only patients with residual deficit are tested. Secondly, the reorganisation process which occurs after stroke might obscure the initial importance of some brain areas for a specific function. By examining acute patients one can avoid this problem (Hillis, Barker et al. 2000). In the first study of this kind, Hillis et al. (2000) correlated behavioural results with estimated penumbra volume, as defined by examining the mismatch between acute (within 24 hours) DWI images and subacute (after 3 days) perfusion images. The authors also described areas in which reperfusion brought about behavioural improvement. A subsequent study found an association between cortical hypoperfusion and specific language impairment (Hillis, Wityk et al. 2002). Other studies which looked at the relation between behaviour and acute hypoperfusion by employing an ROI analysis found an association between apraxia of speech and hypoperfusion of Broca's area (Hillis, Work et al. 2004), auditory comprehension impairment and hypoperfusion in BA22 (Hillis, Kleinman et al. 2006), improvement on picture naming and acute hypoperfusion in BA37, 22 and 44/45 (Hillis, Kleinman et al. 2006), and between hypoperfusion in subsets of the left hemispheric language network and specific processes underlying picture naming (DeLeon, Gottesman et al. 2007). Croquelois et al. (2003) defined a specific set of left hemispheric ROIs as potentially relevant for each aphasia subtype. For example, the frontal lobe and basal ganglia were defined as relevant for verbal fluency, and the temporal, frontal and insular cortex for comprehension deficits. An association was found between the number of hypoperfused ROIs (within the predefined set) and overall comprehension, verbal fluency and repetition impairments, but not naming deficit (Croquelois, Wintermark et al. 2003). This method still has one of the main disadvantages of the lesion-symptom mapping techniques, namely, using ROIs. While the results of such perfusion studies can be relevant to the current research, this study focuses on chronic patients and therefore used VLSM. Moreover, VLSM allows a wider anatomical analysis, which is not focused on predetermined ROIs.

1.2 Methodological Aspects of VLSM

Validation of VLSM, and further analyses on the same or similar cohort of patients, was done using various language tests (Bates, Wilson et al. 2003; Dronkers, Wilkins et al. 2004; Saygin, Wilson et al. 2004; Wilson and Saygin 2004; Baldo, Schwartz et al. 2006; Baldo and Dronkers 2007; Borovsky, Saygin et al. 2007; Dick, Saygin et al. 2007). In these studies, the groups of patients have been seen in clinics over many years and the studies were done retrospectively. As a consequence, the data set contains a combination of CT and MRI scans acquired with a variety of parameters. A trained neurologist drew the lesions manually onto a set of 11 axial template slices. The 11 slices were approx 6 mm apart and the pixel size within each was approx 0.5 x 0.5 mm. However, the spatial accuracy in VLSM is determined not only by the voxel size, but also by the experimenter's lesion reconstruction (Stephen Wilson, personal communication, May 2008). The patient pool consisted initially of 101 patients and was reduced with time to 20 patients. Two other independent groups used the technique for studying language impairments, with 16 and 21 chronic patients, respectively (Piras and Marangolo 2007; Richter, Gerwig et al. 2007). The technique was utilised by groups investigating attention (Grandjean, Sander et al. 2008; Molenberghs, Gillebert et al. 2008), executive function (Ploner, Gaymard et al. 2005), motion detection (Saygin 2007) and motor deficits (Schoch, Dimitrova et al. 2006), amongst others. Most studies used the reconstruction technique, while some drew the lesion directly onto T1 or T2 MRI scans. The main difference between the two methods of drawing lesions is that when lesions are drawn directly on the patient's scan, the lesions need to be normalised later. In the reconstruction technique, on the other hand, the lesions are drawn onto a normalised template, and therefore there is no need for normalisation.

In summary, VLSM is an efficient method for mapping behaviour onto lesions in chronic stroke patients. In this study, the neural correlates of inner speech were examined using VLSM, with a long term aim of contributing to diagnosis and prognosis, which in turn can help in defining the patient population which can benefit from therapy using inner speech tasks.

1.3 Neural Correlates of Inner Speech

1.3.1 Studies of Inner Speech

The neural correlates of inner speech have been investigated in functional imaging studies (PET and fMRI). However, some of these studies have methodological limitations, such as lack of control over participants' performance (for example, McCarthy, Blamire et al. 1993; Cuenod, Bookheimer et al. 1995; Ackermann, Wildgruber et al. 1998; Ojemann, Buckner et al. 1998; Hickok, Erhard et al. 2000; Crosson, Sadek et al. 2001; Shergill, Brammer et al. 2002; Okada, Smith et al. 2003; Jancke and Shah 2004; Shergill, Tracy et al. 2006), as well as over overt speech production during the inner speech condition (reviewed in Indefrey and Levelt 2004). The implications of the first limitation are that with no behavioural output it is difficult to know whether participants perform the task using the desired cognitive processes or whether they perform the task at all. For example, McCarthy et al. (1993) found no difference in activation in the left IFG when comparing a covert speech condition and a listening condition. In the absence of any behavioural output it is difficult to determine whether the lack of activation means that the left IFG does not participate in covert speech, or whether participants simply did not attend to the task. Additionally, informative and important data regarding performance (type of response, errors and reaction time) cannot be obtained (Barch, Sabb et al. 1999; Peck, Moore et al. 2004). Many studies did not ensure that participants refrain from producing overt speech when asked to generate only inner speech. The extent of this problem is evident in the conclusion of a recent meta-analysis aimed at identifying areas which are activated only during phonetic encoding and articulation (Indefrey and Levelt 2004). The authors concluded that studies using covert speech did not control for the absence of overt speech and therefore it was difficult to distinguish between brain mechanisms for overt versus inner speech (Indefrey and Levelt 2004). Exceptions are studies which monitored for overt speech by, for example, filming participants' lips (Frings, Dimitrova et al. 2006) or recording potential overt responses during the covert condition (McGuire, Silbersweig et al. 1996).

Obviously, the rhyme and homophone judgement tasks used here can easily overcome the lack of control over participants' performance. By asking participants to make their judgements by using a button press or a similar technology, one can easily make sure that participants are performing the task, and moreover, relevant behavioural data can be obtained. And indeed, rhyme judgement was used by some researchers in the past (Paulesu, Frith et al. 1993; Pugh, Shaywitz et al. 1996; Lurito, Kareken et al. 2000; Poldrack, Temple et al. 2001; Owen, Borowsky et al. 2004; Hoeft, Meyler et al. 2007). Hoeft et al. (2007) found that rhyme judgement created significant activation in posterior parts of the middle and inferior frontal gyrus, the inferior parietal lobule, and lateral regions of the occipital lobe, extending into the inferior temporal lobe. In a study conducted by Lurito et al. (2000) word rhyme judgement task, was compared to a baseline task in which participants performed a similarity judgement on sets of lines. Activation was found in the left hemisphere in the middle frontal gyrus, inferior frontal gyrus, supramarginal gyrus, middle temporal gyrus and fusiform gyrus. In the right hemisphere activation was found in the inferior frontal gyrus. Owen et al. (2004) compared covert rhyme judgement task to an overt homophone reading task, and found significant activation in left pre-central gyrus (BA6), left supramarginal gyrus (BA39 and 40), left inferior parietal lobe (BA40), and left dorsal frontal cortex. Paulesu et al. (1993) found that rhyme judgement activated BA44 as well as motor regions. They suggest that while BA44 is essential for phonological processing, activation in the motor regions is probably related to small laryngeal movements which are not essential for the production of inner speech. Poldrack et al. (2001) compared a rhyme judgement task to a letter case judgement task which acted as baseline. They found only Broca's area (BA44/45) to be activated. Lastly, Pugh et al. (1996) investigated activation associated with non-word rhyme judgement by using an ROI analysis which included the lateral orbital gyrus (BA10 and 47), dorsolateral prefrontal cortex (BA46), inferior frontal gyrus (BA44 and part of 45), superior temporal gyrus (BA22, 38 and 42), MTG (BA21, 37, 39), lateral extrastriate cortex (BA18 and 19), and medial extrastriate cortex. The study showed that frontal areas (lateral orbital, dorsolateral prefrontal and inferior frontal gyri) are specific to phonological processing, while the temporal areas (superior and middle temporal gyri) are involved in phonological processing among other functions.

In summary, in previous functional imaging studies the rhyme judgement task most commonly activates the left IFG and inferior parietal lobe, including the supramarginal gyrus. These studies provide vital information for understanding inner speech, but in order to understand the difference between inner and overt speech, a direct comparison between the two needs to be made.

1.3.2 Studies of Inner and Overt Speech

In order to characterise the relation between inner and overt speech, some brain activation studies compared overt and covert speech directly. Many imaging studies avoid using overt speech because of the technical difficulties that accompany this paradigm, such as the fact that speaking creates movement of the mouth and jaw, which in turn, can produce artefacts as well as misleading activation (Barch, Sabb et al. 1999; Palmer, Rosen et al. 2001; Peck, Moore et al. 2004), or mask significant activation (Birn, Bandettini et al. 1999). The second potential obstacle in performing an fMRI study with overt responses is recording verbal responses in the presence of the scanner noise (Barch, Sabb et al. 1999; Peck, Moore et al. 2004). However, such obstacles can be dealt with (Barch, Sabb et al. 1999; Birn, Bandettini et al. 1999; Palmer, Rosen et al. 2001; Huang, Carr et al. 2002; Peck, Moore et al. 2004; Cusack, Cumming et al. 2005; Shuster and Lemieux 2005; Basho, Palmer et al. 2007), and the studies described below successfully scanned healthy volunteers while they performed various overt and inner speech tasks. Only studies which compared inner speech to overt speech are reviewed below, and unless stated otherwise, the studies used fMRI. Comparing overt and covert responses directly, Basho et al. (2007) found that a covert response to a semantic fluency task produced significantly greater activation in the left middle temporal gyrus (BA21), left superior frontal gyrus (BA6), right cingulate gyrus (BA32), right superior frontal gyrus (BA11), right inferior and superior parietal lobe (BA40 and 7) and the left parahippocampal gyrus (BA35/36). Interestingly, looking for areas of greater activation for the overt speech task (overt>inner), no significant activations were found. The authors attribute some of this activation to inhibition of overt response and response conflict (producing a word but not saying it aloud). Shuster and Lemieux (2005) used a word repetition task. Inner speech showed higher activation in the left MFG and paracentral lobule, as well as in some right hemispheric regions including the post-central gyrus, two regions in the MTG, the precuneus, and the cerebellum. The authors note that the nature of inner speech is task dependent, and that a repetition task might produce inner speech which is different from spontaneous inner speech, or 'thinking in words'. Huang et al. (2002) conducted an ROI analysis, looking at the bilateral mouth area of the primary motor cortex (M1), an area just inferior to it, and Broca's area. They found that Broca's area showed a task-dependent pattern of activation. While in a letter naming task, activation was greater for overt speech, in a task requiring generating animal names, activation was greater for inner speech. Similarly, Bookheimer et al. (1995) found that silent reading, during a PET scan, showed increased activation in the left IFG, compared to reading aloud. Overt reading, on the other hand, showed increased bilateral superior temporal and left supramarginal activation, compared to covert reading. Huang et al. (2002) argue that this increase of activation during silent speech is related to either phonological processing or to the inhibition of overt response. It is unclear however, why phonological processing should differ between inner and overt speech and the issue is not addressed in the study. Ryding et al. (1996) asked participants to count silently and aloud during a PET scan. The two conditions were not directly compared, but by subtracting the rest baseline from each condition, it was found that inner counting produced activation mainly in left hemispheric areas related to speech perception and motor control (SMA, rolandic motor cortex and Wernicke's area), as well as in the right dorsolateral frontal region, that was suggested to be related to focusing attention on the internal activity. This pattern of activation was distinctly different from that obtained for overt speech production and the author suggests that it reflects an internal feedback loop critical for the production of inner speech. Using EEG (Eulitz, Hauk et al. 2000), no early differences between overt and covert speech were found. However, later the signals differed extensively but motion artefacts made it impossible to analyse this data. Focusing on cerebellar activation, Frings et al. (2006) found that silent verb generation was associated with greater activation in a specific right cerebellar region, when compared with overt reading aloud of the same verbs. It is difficult to determine, however, whether the observed activation is related to the difference between inner and overt speech, or to differences in task demands (i.e. verb generation versus reading aloud). It is important to note that these studies also have some of the caveats mentioned above, namely: they failed to control for performance on the inner speech condition and did not ensure that participants refrain from producing overt speech.

In summary, these studies show that vast regions in the brain are activated during various inner speech tasks, when compared to overt speech production, in both hemispheres as well as in the cerebellum. In general, the results of the various studies diverge significantly.

In conclusion, the studies of inner speech alone produce replicable data regarding inner speech but the relation between inner and overt speech is not explored in them. Other studies reviewed here made direct comparison between inner and overt speech but used tasks which do not give information about the participants' performance. It therefore becomes apparent that there is a clear gap in the data available in the literature today. The purpose of the current study was therefore to further our understanding of the neural mechanisms underlying inner speech and its relation to overt speech, while controlling for participants' performance during inner speech generation. In aphasia, lesion analysis together with detailed behavioural testing, can give information regarding the neural correlates of inner speech which cannot be easily obtained by using fMRI. Specifically, lesion analysis can define the areas which are *critical* for, rather than only *contributing* to, the production of inner speech. Lesion analysis also avoids the difficulties in using overt and covert speech in functional imaging studies, mentioned above. As in the behavioural study, rhyme and homophone judgement tasks were used to assess inner speech, and a reading aloud task was used to assess overt speech, with words used in both cases. Based on previous studies, it was hypothesised that one or both of the most commonly activated regions; the left IFG and the supramarginal gyrus, will prove to be essential for inner speech.

2. Methods and Materials

2.1 Participants

After completion of the behavioural study, eligible patients underwent structural brain scans. Out of the entire cohort of patients who participated in the behavioural study (n=29, Chapter 3), 21 were eligible to have an MRI scan (14M/7F; age range: 21-81; mean age: 64±15; mean number of years of education: 12±3; mean time since last stroke: 27±21 months). The other eight patients were excluded due to aneurysm clips (n=3), other metallic clips in the body (n=1), development of vascular dementia (n=1) and claustrophobia (n=3). Further information, including time between last stroke and cognitive-behavioural assessment, time between cognitive-behavioural assessment and head scan, and lesion volume, are available in table 1, Appendix 2. Appendix 2 also specifies which of the patients who participated in the behavioural study, also participated in this VLSM study. The study was approved by the Cambridge Research Ethics Committee and all participants read an information sheet and gave written consent.

2.2 Procedure

2.2.1 Behavioural Testing

Acquisition of behavioural data is described in chapter 3. This experiment analysed three main variables: overt language production, language comprehension and production of inner speech. Analyses of language production and comprehension were used as controls. Two measurement of overt language production were used: reading words aloud and phonological fluency. The language comprehension scores used were those obtained from the auditory sentence and word comprehension tasks, in which participants were asked to match an auditory sentence/word to a picture, and the reading comprehension tasks, in which participants were asked to match a written sentence/word to a picture. Finally, the word homophone and the rhyme judgement tasks were used as measurements of inner speech. Scores of the various language tests, other than the inner speech tasks, were taken from the CAT (Swinburn, Porter et al. 2004), which was completed by all patients.

2.2.2 Imaging Data Acquisition

Structural MR imaging was mainly performed using a 3T Siemens Allegra (Erlangen, Germany) MRI scanner at the Wolfson Brain Imaging Centre, Cambridge. Four patients could not undergo a 3T MRI scan due to having medical devices which were not 3T compatible, including stents (n=2) and PFO devices (n=2). These four patients were scanned using a 1.5T MRI Siemens scanner (Erlangen, Germany) at Addenbrooke's Hospital, Cambridge. Anatomical scans included a Proton Density (PD) and T2-weighted scans (TR: 4.6 sec, TE: 12 msec for PD, 104 msec for T2, FOV: 168x224 mm, matrix: 240x320, sagittal plane; slice thickness: 5 mm; 27 slices), magnetization-prepared rapid-acquisition gradient echo (MPRAGE) scan (TR: 2.3 sec, TE: 2.98 msec, FOV: 240x256 mm, sagittal plane; slice thickness: 1 mm; 176 slices) and axial-fluid-attenuated inversion recovery (FLAIR: TR: 7.84 sec, TE: 95 msec, FOV: 256x320 mm, axial plane; slice thickness: 4 mm; 27 slices).

2.2.3 Data Analysis

Lesions were defined using the Regions of Interest facility in Analyze software (Mayo Biomedical Imaging Resource, Mayo Clinic, MN). The author (SG) drew the lesions manually. Lesions were drawn on T2 scans, while consulting other sequences (T1, PD, and FLAIR). After drawing on a slice, the drawn object was copied to the next slice (above or below) and adjustments were made. Lesion definitions resulted in images that had the value of 1 for lesioned voxels and 0 for normal tissue. This technique of lesion masking was validated previously and has been shown to be reliable for unilateral lesions, even if the lesion is as big as the entire hemisphere (Brett, Leff et al. 2001). The masks drawn were validated by a trained neurologist (EAW) who was blinded to the patients' diagnosis. Normalisation was done using MRIcron (Chris Rorden's MRIcron, 2007) and the Statistical Parametric Mapping software (SPM5, Wellcome Department of Cognitive Neurology, UCL) implemented in the MATLAB (2006b) environment. Lesion analysis was done using VLSM version 1.6 (Bates, Wilson et al. 2003), in which for each voxel, the behavioural scores of those patients with a lesion in this voxel, and those without, are compared using a t-test. This procedure is then repeated for every voxel in which at least one patient has a lesion. Correction for multiple comparisons is done using the False Discovery Rate (FDR) at a corrected threshold of $p<0.05$. Other variables examined were age, lesion volume and time post-stroke (months).

To tease apart the different components of each of the inner speech tasks, measurements of overt speech production and verbal working memory were included. Table 4.1 specifies the cognitive processes involved in each task. By adding covariates to the analyses, specific cognitive components can be isolated.

Table 4.1. Cognitive sub-processes involved in the inner speech tasks.

Test	Visual word processing	Grapheme to phoneme translation	Inner speech (phonological output to phonological input)	Phonetic coding and articulation[12]	Verbal working memory
Homophone judgement	Y	Y	Y	N	N
Rhyme judgement	Y	Y	Y	N	Y
Reading word aloud	Y	Y	N	Y	N
Sentence repetition	N	N	N	Y	Y

The comparisons of interest in this study were:

1. [Homophone judgement > reading words aloud] examines inner speech alone.
2. [Rhyme judgement > reading words aloud] examines inner speech and working memory.
3. [Rhyme judgement > sentence repetition] removes the working memory component in the rhyme judgement task.

Statistical maps are superimposed on a normalised single subject brain template.

3. Results

17 patients were scanned on the 3T scanner at the Wolfson Brain Imaging Centre and 4 patients were scanned using the 1.5T MRI scanner at Addenbrooke's Hospital. There were technical problems with the scan of one patient (DR, scanned at 1.5T) and the data of this patient was therefore excluded from all analyses. Correlations between age, time since stroke or lesion volume, and performance on the inner speech tasks were tested using Kendall's Tau (significance level was determined at $p<0.005$, after Bonferroni correction for multiple comparisons). Age and time since stroke did not significantly correlate with performance on the inner speech tasks ($p>0.005$ for all). Lesion volume significantly correlated with performance on the

[12] Phonetic coding and subvocal articulation can accompany inner speech production but it is not a necessary component.

rhyming task (Kendall's Tau=-0.68, p<0.001) and on the homophone task (Kendall's Tau=-0.49, p=0.004).

All areas listed in the lesion analyses below were in the left hemisphere. Colour maps are presented in radiological space, with the right hemisphere on the left side, and left hemisphere on the right side. Colour maps correspond to the p values associated with the calculated t-statistics. The actual presented values are –log(p), and the scales run from the thresholded alpha (p<0.05 corrected using the FDR), to the maximum value in the image.

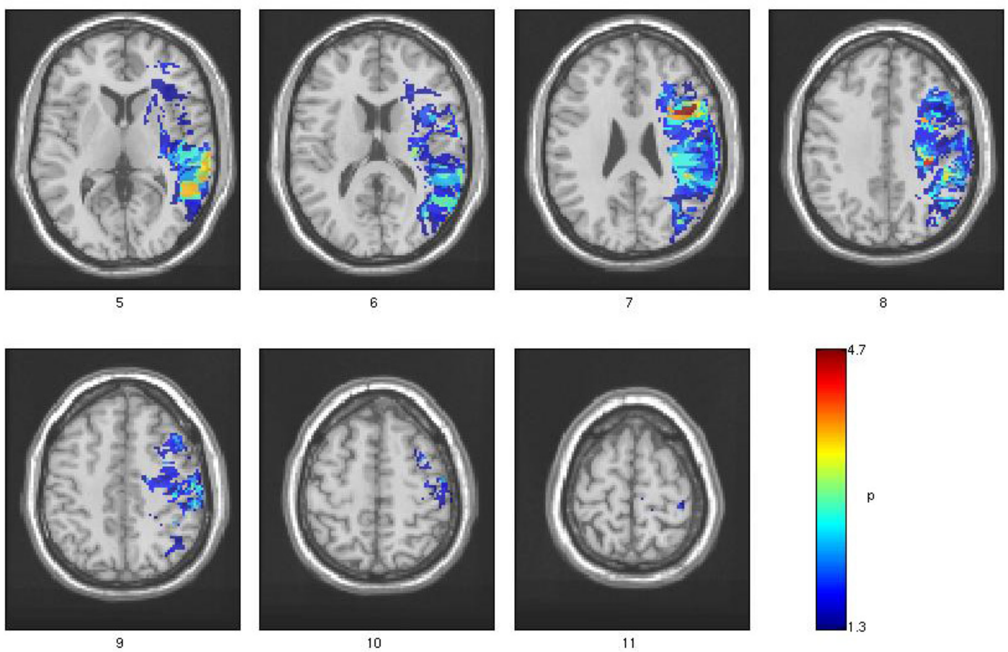
3.1 Neural Correlates of Language Production and Language Comprehension

As control tasks, the anatomical correlates of speech production and speech comprehension were examined. Scores on the **phonological fluency** task were significantly correlated with lesions in the left IFG (pars opercularis, BA44). The language production task most closely related to the inner speech tasks used in this study is **reading words aloud**. Scores on this task were significantly correlated with lesions in the posterior part of the MTG. In addition, various comprehension tasks were examined. Performance on the **auditory sentence-picture matching** task was correlated with a large cluster in the middle section of the MTG. **Auditory comprehension of words** (auditory word-picture matching) was correlated with lesions in the posterior left MTG. **Reading comprehension of sentences** was correlated with two clusters: one in the inferior temporal lobe, extending into the MTG, and the second in the ventral part of the superior frontal gyrus, extending into the MFG, while **reading comprehension of words** was correlated with the posterior inferior temporal gyrus including the fusiform gyrus.

3.2 Neural Correlates of Inner Speech

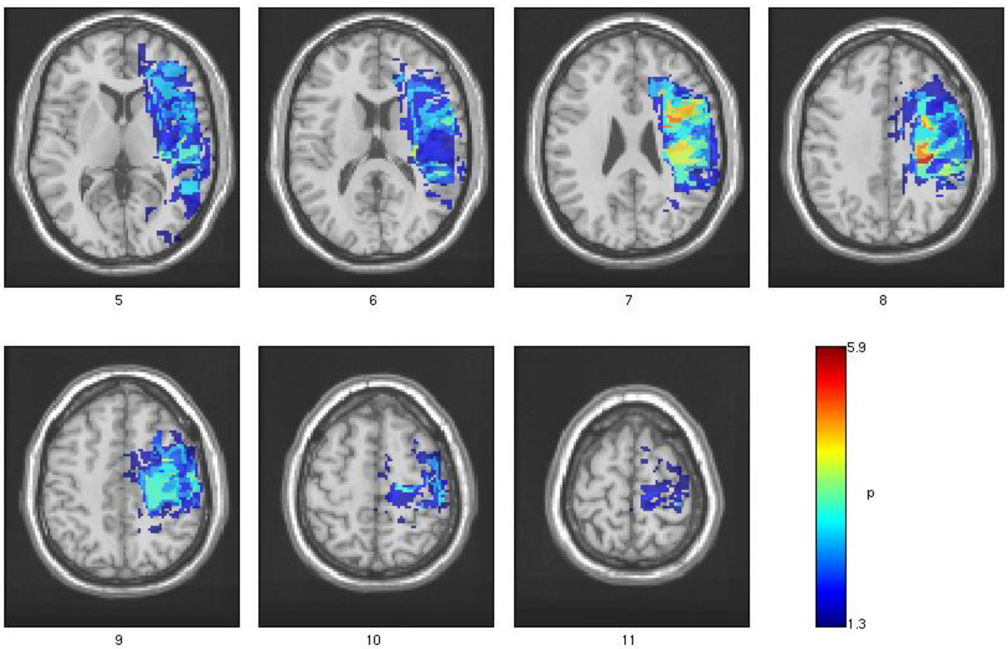
Next, the neural correlates of the inner speech tasks were examined. Performance on the **homophone task** was significantly correlated with lesions in the IFG, pars opercularis (BA44) and the frontal operculum, as well as the posterior part of the MTG (BA21, extending inferiorly into BA37; figure 4.1, significant corrected –log(p)=3.4). Adding age or time since stroke as a covariate did not change these results. However, when adding lesion volume as a covariate, the cluster in the IFG was significant only with uncorrected p<0.001 and the cluster in the MTG only approached this level of significance.

Figure 4.1. VLSM map showing p values of all voxels for homophone judgement. All p values are in –log10. Significant corrected threshold is –log10(p)=3.4



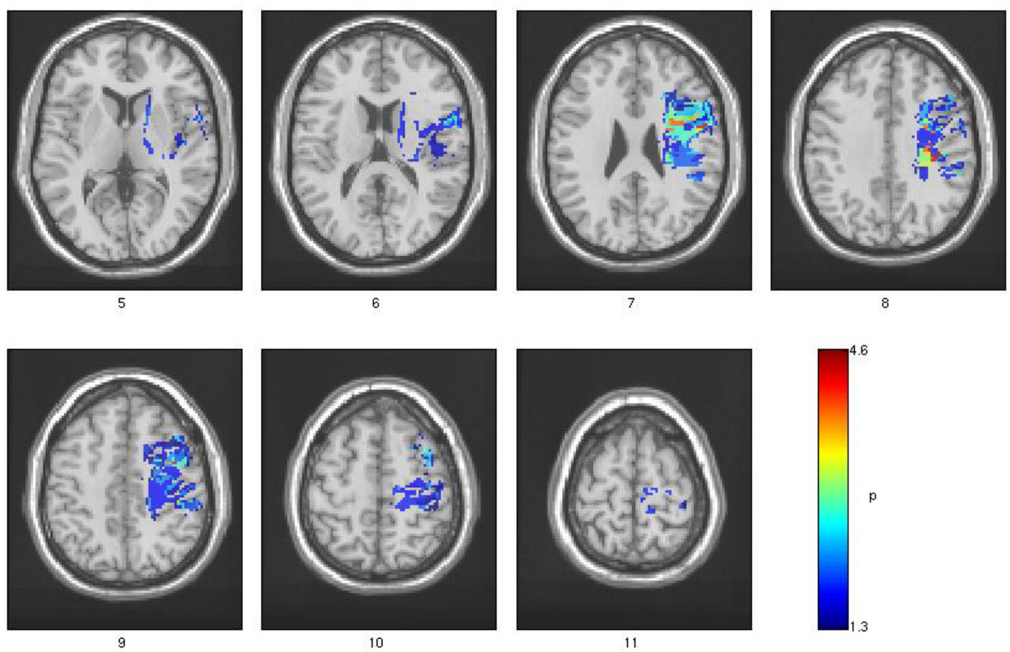
Significant correlates of performance on the **rhyming task** included the left pars opercularis as well, but also the anterior part of the supramarginal gyrus (BA40), extending into the white matter (WM) medial to it (figure 4.2, significant corrected –log(p)=1.95). As with the homophone judgement, adding age or time since stroke as a covariate, did not change these results. Again, adding lesion volume as a covariate resulted in no significant results at the strict threshold level of p<0.05 corrected with FDR. However, the two clusters, in the IFG and in the supramarginal gyrus, were significant at the uncorrected threshold level of p<0.001.

Figure 4.2. VLSM map showing p values of all voxels for rhyme judgement. All p values are in –log10. Significant corrected threshold is –log10(p)=1.95.



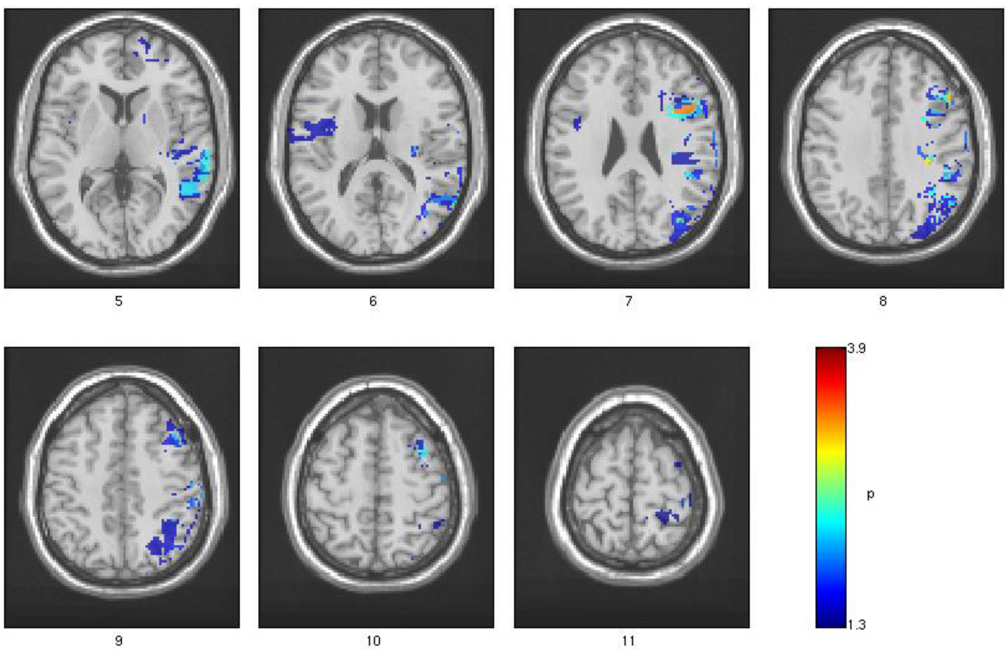
When factoring out the influence of verbal working memory, by adding the scores from the sentence repetition task as a covariate in the analysis, it was found that the rhyme judgement significantly correlated with a small region in the IFG, pars opercularis, extending posteriorly and medially into the WM, as well as a small cluster of WM superior and medial to the supramarginal gyrus. The supramarginal gyrus itself was not significant (figure 4.3, significant corrected –log(p)=3.14). However, a similar analysis in which word repetition was factored out showed that in this case as well the supramarginal gyrus was no longer significant, while lesion to the left IFG was significantly correlated with performance (significant corrected –log(p)= 2.95).

Figure 4.3. VLSM map showing p values of all voxels for rhyme judgement, factoring out verbal working memory. All p values are in –log10. Significant corrected threshold is –log10(p)=3.14.



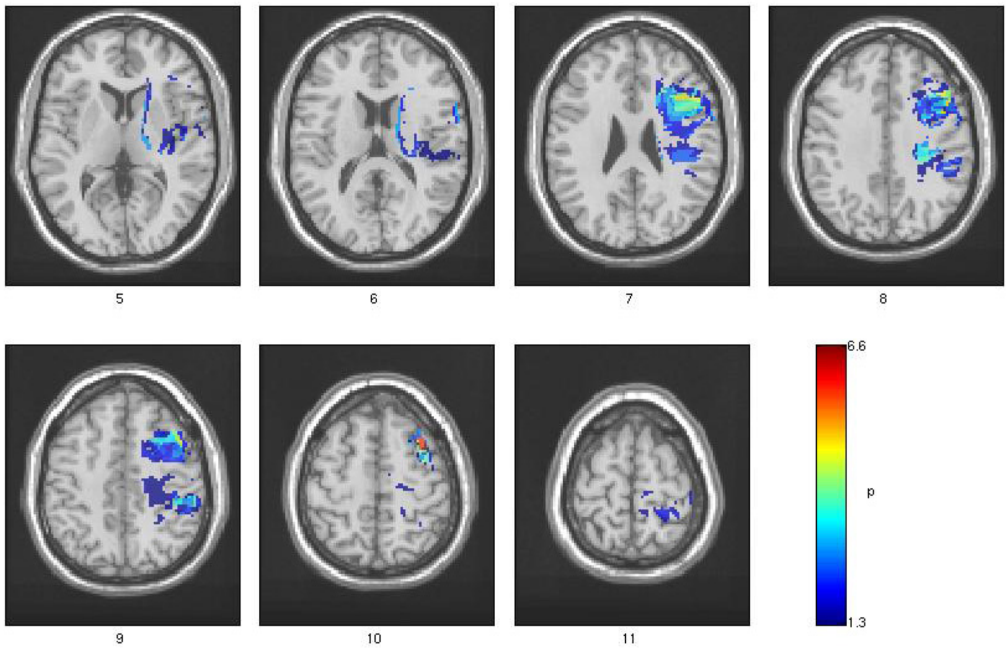
Another question addressed was whether there are areas affecting inner speech abilities over and above overt speech production. To answer this, the neural correlates of the inner speech tasks were identified, while factoring out the influence of overt speech, by including the scores from the word-reading-aloud task as a covariate in the analysis. It was found that homophone judgement was correlated with a small cluster in the IFG, pars opercularis (BA44). This region was significant only with a threshold of uncorrected p<0.001 (figure 4.4, significant uncorrected –log(p)=3.0).

Figure 4.4 VLSM map showing p values of all voxels for homophone judgement, factoring out overt speech production. All p values are in –log10. Significant uncorrected threshold is –log10(p)=3.0.



Rhyme judgement was also correlated with a cluster in the IFG, which here included both the pars opercularis (BA44) and the pars triangularis (BA45), extending posteriorly into the pre-central gyrus. Another cluster was found in the supramarginal gyrus (BA40) and the WM medial to it (figure 4.5, significant corrected –log(p)=2.89)

Figure 4.5. VLSM map showing p values of all voxels for rhyme judgement, factoring out overt speech production. All p values are in –log10. Significant corrected threshold is –log10(p)=2.89.



4. Discussion

This study looked at the neural correlates of inner speech by testing aphasia patients on inner and overt speech abilities and correlating these behavioural results with patients' lesions.

4.1 Lesion Analysis of Language Production and Comprehension

The neural correlates of speech comprehension and speech production were examined, with the aim of testing the validity of the materials used in this study. These functions were studied in details by others (Bates, Wilson et al. 2003; Dronkers, Wilkins et al. 2004; Saygin, Wilson et al. 2004; Baldo, Schwartz et al. 2006; Baldo and Dronkers 2007; Borovsky, Saygin et al. 2007; Piras and Marangolo 2007; Richter, Gerwig et al. 2007). This study found that reading aloud was correlated with lesions to the posterior part of the MTG, an area known to be related to phonological retrieval in speech production (Indefrey and Levelt 2004). This is consistent with previous studies showing that picture naming (Dronkers, Wilkins et al. 2004; Piras and Marangolo 2007) and semantic variety (a measure of spontaneous speech production, Borovsky, Saygin et al. 2007) were also correlated with lesions in this region. Phonological fluency was significantly correlated with lesions to the IFG (pars opercularis, BA44), a finding which replicates results by some (Baldo, Schwartz et al. 2006; Borovsky, Saygin et al. 2007) but not all (Bates, Wilson et al. 2003).

Auditory comprehension of words and sentences was mostly related to lesions in the middle and posterior part of the MTG, respectively, a finding which was reported previously (Bates, Wilson et al. 2003; Dronkers, Wilkins et al. 2004; Baldo and Dronkers 2007). This finding is consistent with current studies of language processing which suggest that phonological-semantic mapping occurs in the posterior MTG and the inferior temporal sulcus (Price 2000; Friederici 2002; Martin 2003; Hickok and Poeppel 2007; Saur, Schelter et al. 2009). Lastly, reading comprehension was found to be associated mainly with the inferior temporal lobe, an area which overlaps with the ventral reading stream (reviewed in Pugh, Mencl et al. 2000; Schlaggar and McCandliss 2007), and includes the left fusiform gyrus. Saygin et al. (2004) did not find involvement of the inferior temporal gyrus in reading comprehension. However, the task used in their study was not a simple sentence/word-picture-matching task, but rather a task which required higher semantic processing. This difference in task-demands might be responsible for the different results.

In summary, the neural correlates of language production and comprehension emerging from this study are consistent with current models and other lesion studies of language, and therefore confirm the reliability of the technique in this cohort of patients. It is therefore possible to now turn to discuss the neural correlates of inner speech.

4.2 Neural Correlates of Inner Speech

This study is the first to look specifically at the relation between inner speech, overt speech and lesion site in post-stroke aphasia. It found that inner speech was consistently correlated with lesions involving the left IFG, especially the pars opercularis (BA44). Rhyme judgement was correlated also with lesions involving the supramarginal gyrus. These areas were statistically significant even when factoring out the influence of overt speech production, although the supramarginal gyrus was not significant when factoring out the influence of verbal working memory. Therefore, the main hypothesis, that the left IFG and the left supramarginal gyrus are related to inner speech, was confirmed. This result is consistent with functional imaging studies of inner speech, especially of rhyme judgement (Paulesu, Frith et al. 1993; Pugh, Shaywitz et al. 1996; Lurito, Kareken et al. 2000; Poldrack, Temple et al. 2001; Owen, Borowsky et al. 2004; Hoeft, Meyler et al. 2007). On the other hand, the studies comparing inner and overt speech show more complex results. These studies used various types of tasks and the nature of these tasks might be key to the interpretation of the results. The rhyme and homophone judgement require a level of active 'use' of inner speech, in a way that one has to monitor, or listen to, his/her own inner speech in order to successfully perform the task. As discussed in the previous chapter, this requires transferring output phonology to the input phonological store. This might be the case also in the semantic fluency task (used in Basho, Palmer et al. 2007) and when generating names of animal (used in Huang, Carr et al. 2002), where the participant needs to keep track of the words already produced. In the case of word repetition (used in Shuster and Lemieux 2005), letter naming (used in Huang, Carr et al. 2002) and silent reading (used in Bookheimer, Zeffiro et al. 1995), and even more so in counting (used in Ryding, Bradvik et al. 1996), which is an automatic process, such monitoring of inner speech is less needed. Using the terminology of Vigliocco and Hartsuiker (2002), it is suggested that the tasks used in this study, along with the semantic fluency task and generating animal names, most likely require 'conscious inner speech'. In contrast, the other tasks require only 'unconscious inner speech'. The neural correlates associated with these two types of inner speech can potentially be very different. Re-examining the results of the study by Huang et al. (2002) shows that generating animal names, a task which requires a more conscious type of inner speech, produced high left IFG activation, while naming letters, which requires a less conscious type of inner speech, did not produce significant left IFG activation. Therefore, it is suggested that the left IFG is more closely related to the former type of inner speech. This is consistent with the data from this study showing that the left IFG, and particularly the pars opercualris, are involved in inner speech. Two results from previous studies seem to contradict the explanation given here. Bookheimer et al. (1995) used a silent reading task, but still had higher left IFG activation for silent speech. Basho et al. (2007), on the other hand, used the semantic fluency task, but did not record left IFG nor supramarginal activation. While these results seem to contradict the explanation given above, it should be noted that in both cases there was no record or indication as to what participants actually did during the scan, nor of their level of performance, cooperation or engagement in the task. It is therefore difficult to directly compare the result of the various studies. Lastly, an interesting study evaluated neural correlates associated with verbal transformations (Sato, Baciu et al. 2004). 'Verbal transformation' refers to the phenomena where a word is repeated rapidly, and after a while a new percept 'pops-out', and the participant starts perceiving a word that is different from the one perceived initially. For example, if the word 'life' is repeated rapidly, it might after some time sound like 'fly'. Sato et al. compared two conditions: in the first condition participants were asked to simply repeat the word, while in the second they were asked to pay attention to the moment in which a verbal transformation occurs. In this way, the authors created two conditions: unconscious inner speech (the former) and conscious inner speech (the latter), by using the same stimuli. Comparing the two conditions directly (attending to the verbal transformation>repetition), they have found that conscious inner speech was significantly correlated with activation in the anterior part of the right cingulate gyrus, left IFG, left supramarginal gyrus, bilateral cerebellum and left STG, a system which includes the two areas found in this study - the left IFG and the supramarginal gyrus.

It should be clarified that it is not proposed that a clear-cut dissociation between conscious and unconscious inner speech can be made. Rather, it is suggested that different tasks require different levels of involvement of the input phonological store, and that this, in turn, influences the involvement of the left IFG. Tasks such as letter naming or counting can be considered as relatively automatic, unconscious, inner speech tasks, and thus require little involvement of the left IFG, while tasks such as rhyme and homophone judgement require more conscious inner speech, and accordingly more involvement of the left IFG. The involvement of the left IFG, and particularly the pars opercularis (BA44) in phonological processing is of course not new and have been shown by many studies in the past (for example Mummery, Patterson et al. 1996; McDermott, Petersen et al. 2003; Burton, LoCasto et al. 2005; Price, Devlin et al. 2005). Moreover, the results from this study, showing that performance on the phonological fluency task is associated with lesions to this area as well, give further support to this idea.

In their case study, Howard and Franklin (1990) suggested that homophone and rhyme judgements also differ in the demands they put on the output phonological store. They argued that rhyme judgement puts more demand on the output phonological store than homophone judgement. This leads to the discussion of the role of the supramarginal gyrus, which was found to be involved in the rhyme judgement task, but not in the homophone judgement task.

Firstly, it should be noted that when factoring out verbal working memory, the involvement of the supramarginal gyrus was eliminated. This suggests that the role of the supramarginal gyrus in the rhyme judgement has more to do with the phonological loop, or working memory. And indeed, studies suggest that the phonological short term store, which is the memory component of the phonological loop (the other component, the articulatory rehearsal process, is more analogous to subvocal speech), is located in the inferior parietal lobe, in BA40 (Paulesu, Frith et al. 1993; Baddeley 2003; Mueller and Knight 2006).

An alternative explanation is that the supramarginal gyrus was factored out when including the sentence repetition task, because of its involvement in repetition, over and above memory. That is, it is suggested that the supramarginal gyrus was factored out due to the repetition component of the task chosen (i.e. sentence repetition), rather than the working memory component *per se*. Wise et al. (2001) have suggested that the junction between the posterior supratemporal region and the inferior parietal lobe acts as a centre for binding speech perception and speech production, or lexical recall. This is relevant for tasks as repetition and auditory comprehension, but also for inner speech production. And indeed, the involvement of the supramarginal gyrus (Quigg, Geldmacher et al. 2006) together with the superior STG (Anderson, Gilmore et al. 1999) in repetition, has been demonstrated previously. Results from this study give further support to this alternative explanation: when factoring out the scores from the word repetition task, in which the memory load is minimal or at least substantially reduced in comparison with the sentence repetition task, it was found that the supramarginal gyrus was not significantly correlated with rhyme judgement. It is acknowledged, however, that based on the data from this study it is impossible to distinguish conclusively between these two alternative explanations.

The lesion analysis also identified some areas of white matter which seem to be part of the dorsal route of language (medial to the supramarginal gyrus and medial and superior to the IFG). These areas were associated with inner speech over and above working memory. Clearly, an analysis looking specifically at white matter (for example, by using Diffusion Tensor Imaging - DTI) could shed more light on the exact anatomy and function of these white matter tracts. However, these results suggest that the dorsal route for language is involved in inner speech. DTI studies demonstrated that the supramarginal gyrus and BA44 are connected via a direct connection (Catani, Jones et al. 2005; Frey, Campbell et al. 2008), as well as via indirect connections through the posterior STG (Parker, Luzzi et al. 2005; reviewed in Friederici 2009). Recently, it was shown that the pars opercularis shows functional connectivity to the supramarginal gyrus during resting state, and it was suggested that the two regions function together as a phonological processing system (Xiang, Fonteijn et al. 2010). The role of the dorsal route in supporting inner speech might be to transfer the output phonological code from anterior areas such as the IFG, to posterior regions, where it is further processed. It should be noted that previous studies emphasised a specific functional directionality of the dorsal route, in which processing advances from posterior to anterior regions (and not vice versa), supporting repetition, language acquisition, and monitoring of overt speech (Catani, Jones et al. 2005; Hickok and Poeppel 2007; Saur, Kreher et al. 2008; Friederici 2009; Agosta, Henry et al. 2010), although Catani et al. emphasise that in DTI studies '…the terms *origin* and *termination* are arbitrary because tractography is blind to fiber direction' (Catani, Jones et al. 2005, p.9). A study of the macaque brain showed that area 44, which is homologues to BA44 in the human brain, receives input from area PFG of the inferior parietal lobule, which is homologues to the caudal part of the supramarginal gyrus in humans (Petrides and Pandya 2009). Matsumoto et al. (2004) used electrodes to directly stimulate anterior and posterior cortical regions and to record evoked potentials in humans. Language areas were defined as those which, when stimulated, impair sentence reading in the individual patient. Anterior regions included Broca's area or adjacent regions and posterior regions included the supramarginal gyrus, the middle and posterior STG and the adjacent MTG. As expected, stimulation of posterior language areas resulted in evoked potentials in anterior language areas, supporting the idea of progress of processing from posterior to anterior. However, stimulation of anterior regions also resulted in evoked potentials in all posterior regions tested, including the supramarginal gyrus, middle and posterior STG and the adjacent MTG. In the study of the macaque's brain described above (Petrides and Pandya 2009), the authors studied fibres which project from posterior to anterior regions, but not vice versa. Since this study or others did not examine whether the pars opercularis sends fibres to the supramarginal gyrus, in the monkey's brain, the possibility that such connection exist cannot be ruled out. In summary, although studies emphasis a propagation of information in the dorsal route from posterior to anterior parts, it is possible that some fibres in this route send information in the other direction, and these might be essential for inner speech production.

4.3 Study Caveats

This study could not unambiguously define the role of the supramarginal gyrus in inner speech, repetition and working memory. Future studies should employ more finely defined behavioural tasks in order to disentangle the various cognitive components of these interacting and overlapping linguistic processes. On top of this specific caveat, VLSM has some general disadvantages. First, VLSM only looks at areas where the lesions are defined. Therefore, one cannot draw conclusions about areas which are not analysed but might be relevant for the specific function. Moreover, a bias in the selection of areas might arise from a few different sources. Firstly, it is clear that the results of every VLSM study are highly dependent on the participating patient cohort. A larger patient cohort, with heterogeneous lesions, could potentially reveal more areas which are crucial for inner speech processing. Secondly, after stroke there is often a distortion of the grey and white matter due to gliosis and atrophy, with changes in the shape and relative size of the ventricles. However, when drawing the lesion, enlarged ventricles are usually not included in the lesion. As a result, periventricular areas which are damaged in many patients may be excluded from the analysis. Thirdly, after stroke, and especially in the chronic stages, remote areas might show Wallerian degeneration. These areas might not be clearly visible on the structural images and therefore will not be included in the lesion definition, although they potentially influence the behaviour under investigation. Another unrelated difficulty with the VLSM technique is that drawing lesions manually is highly time-consuming and potentially somewhat subjective. Lastly, VLSM can only explore anatomical correlates of a specific function in stroke patients. Obviously, the brain of a chronic stroke patient cannot be directly compared to that of a healthy individual, since it underwent anatomical distortions, spontaneous reorganisation and therapy-related reorganisation. All of these influence the anatomy of language at the time of testing.

4.4 Conclusions and Further Research

This study investigated the neural correlates of inner speech using a lesion analysis method. It was found that inner speech is dependent on the left IFG, pars opercularis, and the left supramarginal gyrus, with the two structures having potentially different roles in the production and processing of inner speech. It is suggested that the dorsal route for language is involved in processing inner speech, by transferring information from anterior to posterior regions. It was further acknowledged that VLSM has a few major drawbacks, and some newer techniques try to overcome some of them. One of these is Voxel Based Morphometry (VBM), a techniques used today to correlate structure and function. By being an automated, whole brain analysis, VBM can potentially overcome the main problems of VLSM, including the potential biases in lesion selection and the fact that lesions are defined manually. Moreover, VBM can be used to explore individual differences in healthy populations. The next study, presented in chapter 5, explored the neural correlates of inner speech on the same patient cohort, using VBM. It explores the differences and similarities between the techniques while highlighting the advantages and disadvantages of each method. It was hypothesised that the use of a whole brain analysis can reveal other regions which are essential for inner speech production, which were not uncovered by this study.

Chapter 5

Voxel Based Morphometry Study of Inner Speech

Neither the naked hand nor the understanding left to itself can effect much. It is by instruments and helps that the work is done, which are as much wanted for the understanding as for the hand. And as the instruments of the hand either give motion or guide it, so the instruments of the mind supply either suggestions for the understanding or cautions.

Francis Bacon, The New Organon (1620)

1. Introduction

1.1 The Voxel Based Morphometry Technique

The previous chapter examined the neural correlates of inner speech using VLSM. However, it was recognised that VLSM has limitations, some of which could be overcome by using Voxel Based Morphometry (VBM). Moreover, VBM can be used to explore neural correlates of various functions in the healthy population. This can shed further light on the anatomy of inner speech by providing information which is not available from testing patients alone. This chapter explores the anatomical correlates of inner speech using VBM, aiming to qualitatively compare the techniques and expand on the results obtained in the previous chapter.

VBM aims to identify local differences in tissue composition, after discarding gross anatomical differences between individuals (Ashburner and Friston 2000; Mechelli, Price et al. 2005). VBM can be used to understand the structural differences underlying clinical conditions such as Alzheimer's disease or schizophrenia, as well as correlating behavioural measurements with changes in grey and white matter. This makes it possible to examine individual differences even in healthy populations (Richardson and Price 2009). In the optimized procedure implemented in SPM, normalisation takes place with the unified segmentation–normalisation procedure, in which the brain is segmented into white matter (WM), grey matter (GM) and cerebro-spinal fluid (CSF), and then normalised according to a standard template (Good, Johnsrude et al. 2003). The images are then smoothed. Next, voxel intensities are compared between groups, correlated with behavioural or other measurements, or both, in order to define brain regions that are relevant for the behaviour in question (Ashburner and Friston 2000). It is possible to analyse the entire brain or focus on specific ROIs.

1.2 VBM Studies of Language

1.2.1 Studies of the Healthy Population

Reviewing VBM studies in healthy populations, Richardson and Price (2009) have concluded that results from structural studies correlate well with those from functional studies of language, and at the same time, structural studies can also bring to light novel findings. One of the main limitations of structural studies in healthy populations is that they cannot determine causality. That is, their results cannot answer the question of whether the structural changes observed are the cause or the result of behavioural differences. One explanation for observed changes might be that people with different pre-determined potential, reflected in tissue density in the central nervous system, develop different language abilities. Others might suggest that the acquisition of language functions causes the changes in tissue density observed in the brain. Currently, studies do not distinguish between these two explanations. There are only very few studies looking at language function in healthy participants, using VBM (Richardson and Price 2009). Golestani et al. (2002; 2007) showed that the perception of novel speech sounds is correlated with white matter density in a region just anterior to the parieto-occipital sulcus bilaterally (2002; 2007), as well as in the left Heschel's gyrus, the right pre-central gyrus and the bilateral lingual gyrus (Golestani, Molko et al. 2007). The production of novel speech sounds, on the other hand, was correlated with white matter density in the left insula/prefrontal cortex and in the bilateral inferior parietal cortex (Golestani and Pallier 2007). In a different study of speech production (Grogan, Green et al. 2009), participants were tested with semantic and phonological fluency tasks. Grey matter density in the cerebellum bilaterally was significantly correlated with overall performance. Phonological fluency showed an effect in the head of the caudate nuclei bilaterally and in bilateral pre-SMA regions. Semantic fluency showed no significant effect in the whole-brain analysis (Grogan, Green et al. 2009). These studies show that VBM can be effectively used to explore language processing in healthy population.

1.2.2 Studies of Stroke Patients

Very few studies so far have used VBM to correlate anatomy with behaviour in stroke patients. Probably, because until recently it was difficult to reliably segment and normalise brains which have substantial structural abnormalities, as is the case in stroke. More recently it has been shown that the normalisation procedure implemented in SPM5 onwards can successfully normalise lesioned brains automatically (Crinion, Ashbumer et al. 2007).

An advantage of VBM in stroke studies is that it can be used to address the question of causality. As mentioned earlier, when looking at healthy populations, one cannot determine whether the structural changes are the cause or the effect of any behavioural differences. However, when examining stroke patients, it is clear that the behavioural changes are a consequence of the stroke. Moreover, by selecting a random and matched control group, one can overcome the potential confound of pre-stroke individual differences, which should exist to the same level in the non-clinical population. In the area of language, Leff et al. (2009) have shown that the left STG is associated with performance on the digit span task and on an auditory comprehension task. Rowan et al. (2007) examined the relation between infancy or childhood stroke affecting the basal-ganglia and language impairments. They found that language impairments were not associated with basal ganglia damage, but rather with left hemispheric cortical changes which were too subtle to be detected visually on conventional MRI sequences. Others used VBM in stroke to explore the relation between grey and white matter changes and therapy techniques for motor impairments (Gauthier, Taub et al. 2008), and presence or absence of cognitive impairments (Grau-Olivares, Bartres-Faz et al. 2007; Stebbins, Nyenhuis et al. 2008).

1.3 Age Related Structural Changes in the Brain

The majority of functional brain studies (fMRI or PET) which were reviewed in the previous chapter scanned young healthy volunteers, typically college students, aged around 18-30 years. For this study older volunteers were recruited, who matched the patient group in age and number of years of formal education. It is well documented that ageing brings about changes in brain structure. The most common finding is that although age brings about reductions in grey and white matter throughout the brain, the greatest reduction occurs in the frontal lobes (reviewed in Galluzzi, Beltramello et al. 2008; Grady 2008; Kalpouzos, Chetelat et al. 2009).

Others suggest that changes in grey matter can be seen also in the parietal lobe (reviewed in Galluzzi, Beltramello et al. 2008). It is therefore difficult to compare the results of this study, conducted on older participants, to previous studies, which were mainly conducted on younger volunteers. This study examined two control groups: one of young and the second of older volunteers. By understanding age related differences in the structure of the language system, it would be easier to compare the results from this study of older controls and mostly elderly patients, to studies of younger volunteers, typically found in the literature.

The occurrence of cognitive decline in ageing is well documented and it is generally agreed that decline occurs in specific cognitive domains but not others (Ivnik, Malec et al. 1996; Salthouse 2003; Craik and Bialystok 2006). While studies have shown that memory and executive functions decline with age (Ivnik, Malec et al. 1996; Craik and Bialystok 2006; Grady 2008), language comprehension is considered to be preserved even in very old age (Ivnik, Malec et al. 1996). Language production, on the other hand, shows age-related effects, with word retrieval deteriorating the most (Mitrushina and Satz 1995; Ivnik, Malec et al. 1996; Marien, Mampaey et al. 1998). This is more likely to be related to difficulties in retrieval than to loss of word representations (Burke, Mackay et al. 1991; Heine, Ober et al. 1999; Wierenga, Benjamin et al. 2008). Studies also show that older adults encounter more cases of tip-of-the-tongue (TOT) in both experimental settings (Burke, Mackay et al. 1991; Heine, Ober et al. 1999; James and Burke 2000) and in every day life (Burke, Mackay et al. 1991; Heine, Ober et al. 1999). Despite this common finding, brain studies of language in ageing have focused on various aspects of language comprehension, such as sentence comprehension (Wingfield and Grossman 2006), syntactic processing (Tyler, Shafto et al. 2010), speech recognition (Harris, Dubno et al. 2009), speech perception (Wong, Jin et al. 2009) and word recognition (Brassen, Buchel et al. 2009), rather than on language production. Exceptions are a few current interesting fMRI studies (Wierenga, Benjamin et al. 2008; Galdo-Alvarez, Lindin et al. 2009; Meinzer, Flaisch et al. 2009; Shafto, Stamatakis et al. 2010), which will be discussed in the next chapter. A single VBM study found that with age, occurrence of TOT states increased, and that this correlated with reduction in grey matter in a left hemispheric area covering the left insula, rolandic operculum, Heschel's gyrus and the STG (Shafto, Burke et al. 2007).

1.4 The Current Research

Based on previous functional imaging studies of inner speech, it was hypothesised that grey and white matter density in the areas of the left IFG and the supramarginal gyrus will correlate with inner speech performance, as was found in the VLSM study. Moreover, the previous chapter suggested that the dorsal route for language is responsible for inner speech processing. Since VBM is a whole brain technique which can potentially reveal small differences in tissue composition which are not detectable by the naked eye, it was hypothesised that other regions along the dorsal route of language might be found as well. The dorsal route originates in the dorsal STG, and further involves the temporoparietal junction including inferior parietal regions (Hickok and Poeppel 2007; Saur, Kreher et al. 2008), posterior IFG (Hickok and Poeppel 2007; Saur, Kreher et al. 2008), premotor cortex in the middle frontal gyrus (Wilson, Saygin et al. 2004; Hickok and Poeppel 2007; Saur, Kreher et al. 2008) and the anterior insula (Hickok and Poeppel 2007). Lastly, imaging studies which compared overt and inner speech show that many right hemispheric regions, especially those that are language homologues, are involved in inner speech production. The VLSM study did not examine the involvement of right hemispheric regions since all patients had left MCA stroke. VBM, on the other hand, can examine the involvement of such regions. However, data from lesion studies fail to replicate the importance of these areas in language processing, which suggests that right hemispheric regions are involved, but are not crucial, for inner speech production. It was therefore hypothesised that right hemispheric regions will not be found in this VBM study. Based on findings from previous studies it was also hypothesised that regional age related reduction in grey matter density will be most prominent in the frontal lobe although detected throughout the brain. Lastly, it was hypothesised that within the language system, changes which correlate with age will be significant in frontal regions.

2. Methods and Materials

2.1 Participants

Two groups of healthy volunteers participated in the study: 12 young controls (4M/8F; age range: 21-34, mean age: 24.6±4.5; mean number of years of education:

18.1±2.1) and 19 older controls (8M/11F; age range: 55-71, mean age: 64.1±4.8; mean number of years of education: 15.1±2.9). One young female who was scanned was later excluded from the study, since she revealed that she is ambidextrous after being scanned. All healthy volunteers had no previous history of stroke and no history of neurological, psychiatric or language disorders. They were right handed and native speakers of British English. All of the volunteers also participated in the behavioural study presented in chapter 2. The patient group is identical to the one who participated in the VLSM study, chapter 4. The patient group was comprised of 20 patients who were eligible to have an MRI scan (13M/7F; age range: 21-78; mean age: 64±15; mean number of years of education: 12±3; mean time since last stroke: 27±21 months). The study was approved by the Cambridge Research Ethics Committee and all participants read an information sheet and gave written consent.

2.2 Materials

Materials and the experimental procedure were identical to those used in the VLSM study, presented in chapter 4, and in the behavioural studies, presented in chapters 2 and 3. To summarise, healthy volunteers performed the inner speech battery (rhyme and homophone judgement). Patients performed the inner speech battery and the Comprehensive Aphasia Test (CAT). Other tests performed by patients and healthy volunteers were not used in this study.

2.3 Procedure

All healthy volunteers were scanned in the 3T Siemens Allegra (Erlangen, Germany) MRI scanner at the Wolfson Brain Imaging Centre, Cambridge. Behavioural testing and scanning took place in one session each. For patients, tests were administered in 2-3 sessions, depending on the patient's ability. Test sessions took place either at Addenbrooke's hospital or at the patient's home, according to the patient's preference. The time between the completion of behavioural tests and the scan, for each patient, is available in Appendix 2, table 1.

The procedure was described before, in chapters 2 and 3. In short, in each of the inner speech tasks, patients performed the task on half of the items (one list) using inner speech and half (second list) using overt speech. In both cases, the patient would first read both words in the pair (either internally or overtly, depending on the condition) and then give his/her judgement for the pair. The two conditions were completed separately and successively, and the order of conditions was randomised between patients. Lists were randomised between patients so that half of the patients judged list 1 using inner speech and list 2 using overt speech, and the other half of the patients, vice versa. Patients who cannot read aloud at all (as confirmed beforehand using the CAT) performed the entire task (both lists) using inner speech alone. Healthy volunteers first performed the entire task (both lists) using inner speech alone and immediately afterwards were asked to read aloud all the words.

Scoring of the inner speech task was based on the judgement given to a word pair, with possible answers being correct or incorrect. Hence, every pair judged incorrectly was scored as one error. For the overt reading task, each word was scored separately, so every word read incorrectly was scored as one error. If the participant happened to read both words of a pair incorrectly, this would be scored as two errors. When a self-correction occurred, the word was scored as correctly read aloud.

2.4 Data Analysis

2.4.1 Images Preprocessing

Images were preprocessed using SPM8 (Wellcome Department of Cognitive Neurology, UCL) implemented in the Matlab (2006b) environment. MPRAGE images were automatically normalised and segmented into GM, WM and CSF probability maps, based on the standard Montreal Neurological Institute (MNI) template, using the unified segmentation-normalisation algorithm in SPM8. This procedure combines tissue segmentation, bias correction and spatial normalization in a single unified model (Ashburner and Friston 2005). In the output images, the value in each voxel represents the probability that this voxel belongs to the specific tissue type and not another. Higher values indicate higher tissue density. Normalised unmodulated GM and WM images were visually inspected for quality of the segmentation-normalisation process, and smoothed with a Gaussian kernel of 8 mm full-width-at-half-maximum.

2.4.2 Statistical Analysis

All statistical analyses used the general linear model as implemented in SPM8. Analyses were performed on the smoothed unmodulated images of grey and white matter, since modulated images provide a measure of the original volume, which is difficult to measure around the lesion (Cathy Price, personal communication, December 2009). The grey and white matter images were entered into a multiple regression model. Grey and white matter masks were created using the *apriori* GM and WM templates included in SPM8, with a threshold of voxel intensity equals to, or bigger than, 0.2. A GM mask was used for the analysis of GM and a WM mask was used for the analysis of WM. These masks ensured that ventricles and areas outside the brain are not included in the analysis. Areas reported are those which showed a significant positive effect at a threshold of $p<0.01$ at local maxima, after correction for multiple comparisons using Family Wise Error (FWE, Chumbley and Friston 2009). When the low number of participants significantly reduced the power of the analysis, uncorrected results are reported with a threshold level of $p<0.0001$. All clusters are comprised of a minimum of 20 voxels. The p-values reported are those associated with the local maxima, and not the cluster size. The reason for this is that cluster size reflects, among real statistical differences in question, also the size of the smoothing and some characteristics of the specific brain region. Therefore, the statistics of the cluster size are not a spatially invariant measurement. That is, they vary between different brain areas and therefore cannot be used as an unbiased measurement (Mechelli, Price et al. 2005). To tease apart the different components of each task, measurements of language and memory were included, and analyses were conducted as described in chapter 4. For each of the two inner speech tasks, analysis was conducted in three stages. First, the patient data was analysed separately, and the effects of time since stroke, age and lesion volume were examined by adding each variable separately as a covariate of no interest. Next, the data from the healthy controls was analysed. Finally, the data was collapsed and a final analysis was done by including a group covariate. Statistical parametric maps are superimposed on the normalised single subject brain template included in SPM8.

Since three of the patients were scanned on a different scanner than the rest of the participants, the effect of scanner type was examined. Some of the main analyses were performed again with one of two possible variations: First, analyses were done only with the participants who were scanned on the 3T scanner, which included 17 patients and all of the control participants. Secondly, analyses were carried out for all participants but with adding a covariate describing the scanner type. These variations were performed on the language production and comprehension analyses (phonological fluency, reading word aloud, auditory comprehension of sentences and words, and reading comprehension of sentences and words), the homophone judgement of all participants and the rhyme judgement of all participants.

3. Results

Throughout the result section below, if not specified otherwise, the area reported upon is in the left hemisphere. All images show areas in the left hemisphere only.

3.1 Neural Correlates of Language Production and Comprehension

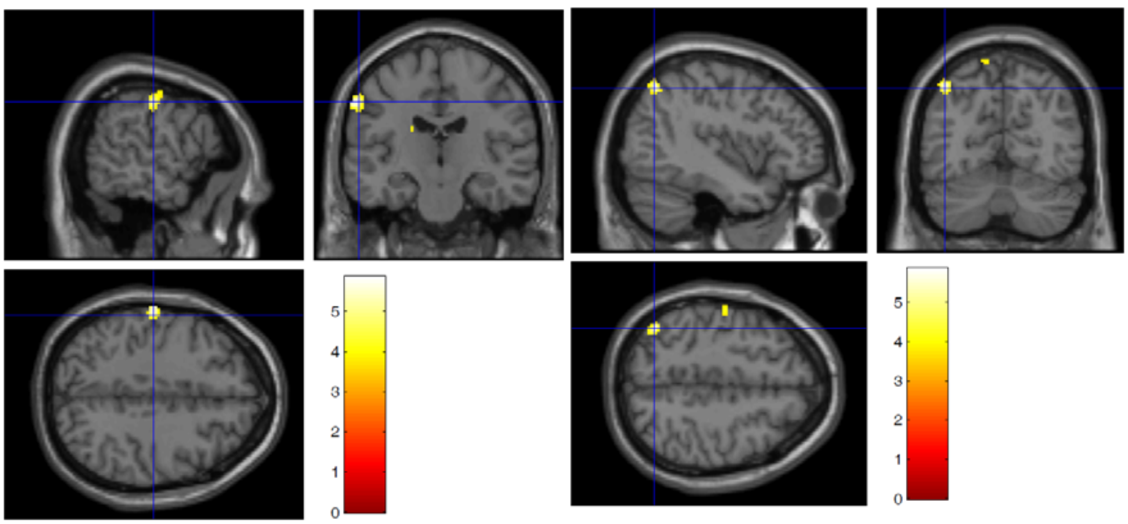
As control tasks, the anatomical correlates of speech production and speech comprehension were examined. All behavioural scores are taken from the CAT, and results are reported with an uncorrected threshold of $p<0.0001$. All patients (n=20) completed the CAT in full. The tasks reported below are the same tasks as those reported for the VLSM study.

Phonological fluency resulted in no significant clusters. Performance on the **reading words aloud** task was significantly correlated with clusters in the left supramarginal gyrus, extending anteriorly and inferiorly into the STG, and further into the post-central gyrus. **Comprehension of spoken sentences** was correlated with GM density in the left anterior and posterior parts of the MTG, while **comprehension of spoken words** was correlated with a small cluster of GM in the left post-central gyrus. **Reading comprehension of sentences** was significantly correlated with GM density in the left and right medial superior frontal gyrus, and the left inferior temporal gyrus (ITG), while **reading comprehension of words** was significantly correlated with two large GM clusters: one in the left fusiform gyrus, posterior inferior temporal lobe and middle section of the MTG, and the other in medial structures of the frontal lobe, including the superior frontal gyrus and the anterior cingulate gyrus.

3.2 Neural Correlates of Inner Speech

Looking at the patient data only (n=19), performance on the **homophone judgement** was correlated with grey matter density in two left hemispheric regions: one in the angular gyrus and the second in the supramarginal gyrus ($p<0.0001$, uncorrected, figure 5.1). Adding time since stroke or age as a covariate did not change the results, while the covariate 'lesion volume', resulted in no significant results.

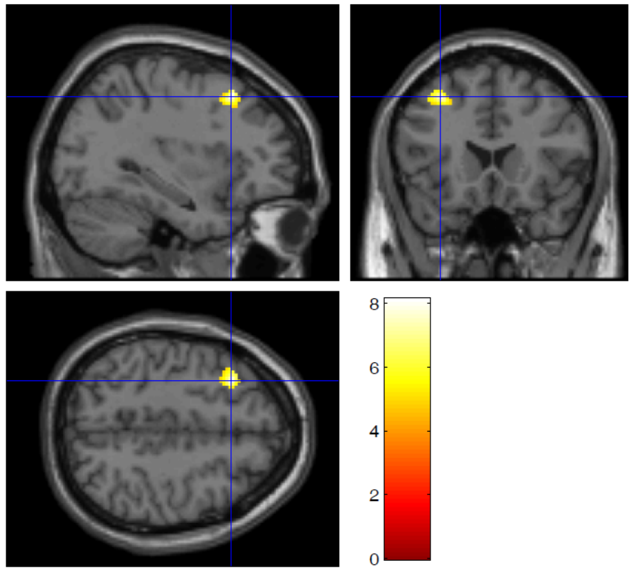
Figure 5.1. VBM maps showing z values of all grey matter voxels for the homophone judgement, in the patient group. All p values are significant at the uncorrected threshold of $p<0.0001$. Areas shown are the left supramarginal gyrus (left) and left angular gyrus (right).



To test whether inner speech differs from overt speech, a variable measuring speech production (scores of reading aloud, taken from the CAT) was included. Both comparisons (inner speech over and above overt speech and vice versa) resulted in no significant results at the specified threshold ($p<0.0001$, uncorrected). Similar results were found when the score for overt speech was the one obtained from the task requiring reading aloud the homophonic words.

The analysis of the WM did not show any significant results, when looking at homophone judgement alone. However, when adding a measurement of overt speech as a covariate, inner speech was significantly ($p<0.0001$, uncorrected) correlated with a cluster of white matter in the MFG, including the frontal eye field (BA8) and the dorsolateral prefrontal cortex (DLPFC; BA9/46, figure 5.2).

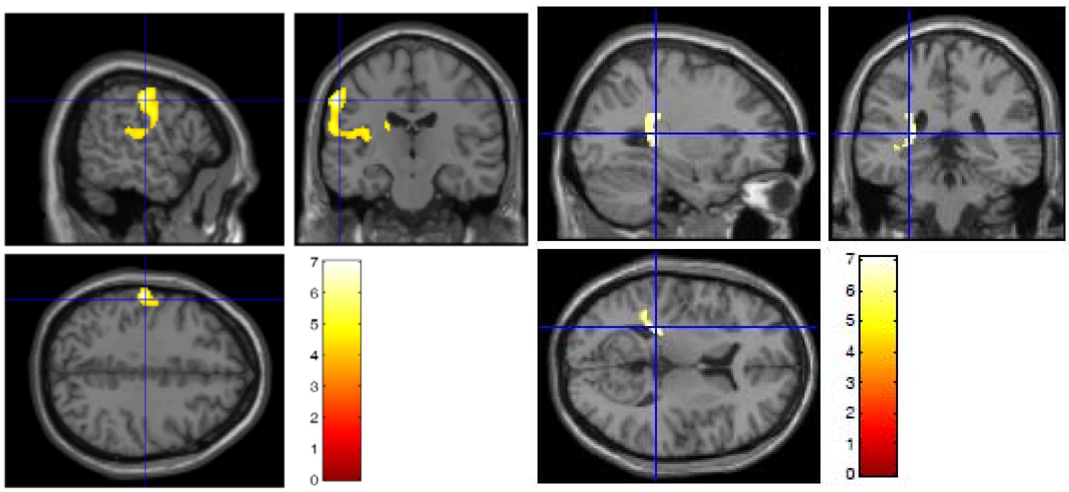
Figure 5.2. VBM maps showing z values of all white matter voxels for the homophone judgement in the patient group, when a measurement of overt speech was included as a covariate. All p values are significant at the uncorrected threshold of p<0.0001. Area shown is in the left MFG.



The correlates of the homophone judgement were then examined across the entire control group (young and older controls together, n=31). Performance on the task did not correlate with any clusters of grey or white matter, at the uncorrected threshold of p<0.0001. There were also no significant clusters when the measurement of overt speech was added as a covariate.

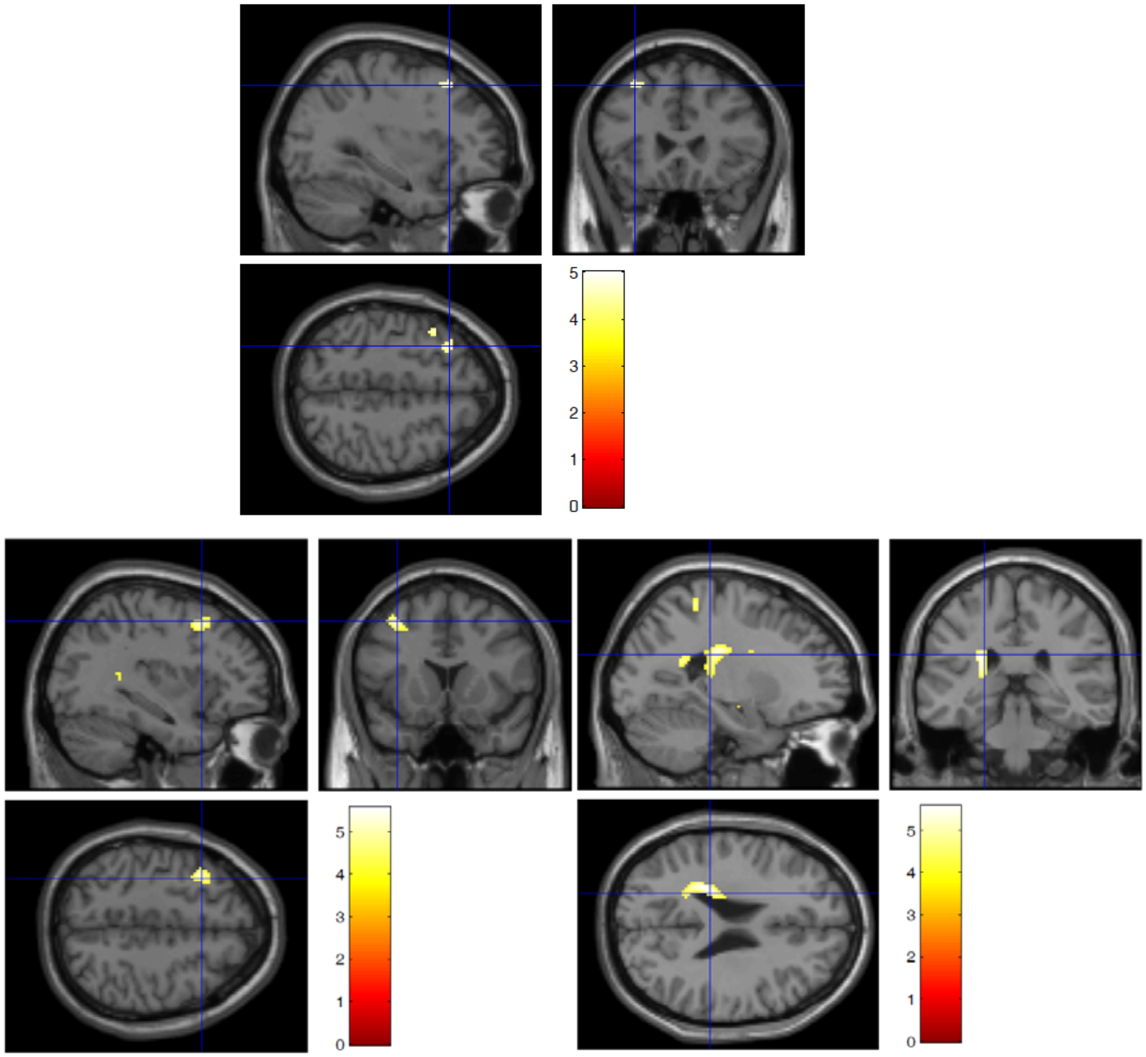
When the data from all participants (patients and controls, n=50) was combined, performance on the homophone task was correlated with grey matter clusters in the insula, extending into the IFG pars opercularis (BA44), and superiorly into the post central gyrus. Another cluster was found in the left angular gyrus (FWE correction, p<0.01). A cluster lining the first and second ventricles and the aqueduct was found as well. This was suspected to be an artefact, arising due to the difference in ventricle size and shape between the patients and controls. When a group covariate (controls *versus* patients) was included in the main design this cluster disappeared and the rest of the results remained similar but only at a more relaxed threshold of p<0.00001, uncorrected (figure 5.3). This suggests that the peri-ventricular blobs are indeed artefacts. Analysing the unmodulated white matter images of all participants, it was found that homophone judgement significantly (FWE correction, p<0.01) correlated with reduced tissue density in a white matter area extending medial to the STG, superiorly towards the supramarginal gyrus and the posterior inferior parietal lobe, and finally, turning anteriorly towards the anterior inferior parietal lobe. When a group variable (controls *versus* patients) was added as a covariate, very similar results were obtained, just at a slightly more relaxed threshold (FWE correction, $p<0.05$). The main difference was that some small clusters around the ventricles disappeared (figure 5.3).

Figure 5.3. VBM maps showing z values of all grey matter (left) and white matter (right) voxels for the homophone judgement, for all participants (group covariate included). All p values are significant at the uncorrected threshold of $p<0.0001$, for the grey matter images, and $p<0.05$ (FWE) for the white matter images.



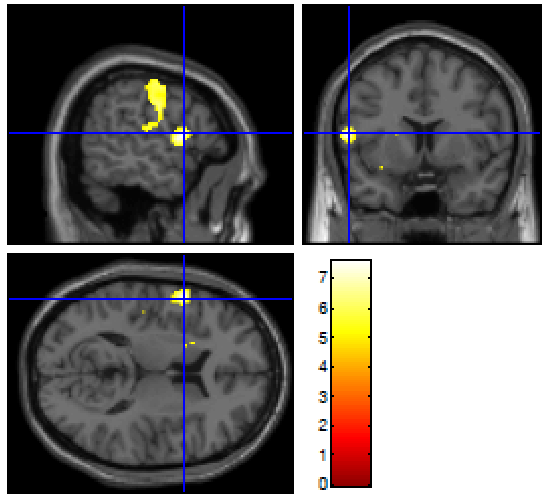
In order to distinguish the cognitive processes related to inner and overt speech, a measurement of overt speech was included. This was the score obtained by reading aloud the words from the homophone judgement task. Inner speech was correlated, over and above overt speech production, with two small grey matter clusters in the MFG; one in the premotor cortex (BA6) and the other in the DLPFC/frontal eye field (BA8/9) ($p<0.0001$, uncorrected), and with two white matter clusters: one medial to the left MFG grey matter clusters (covering both premotor and DLPFC regions, of BA6 and 9/46 respectively, and extending into the superior part of BA44), and the other similar to the one obtained in the analysis of the homophone judgement alone ($p<0.0001$, uncorrected; figure 5.4).

Figure 5.4. VBM maps showing z values of all grey matter (top) and white matter (bottom) voxels for the homophone judgement, for all participants, when a measurement of overt speech was included as a covariate. All p values are significant at the uncorrected threshold of p<0.0001.



The second inner speech task was the **rhyme judgement**. Looking at the data from the patients only (n=17), performance on the rhyme judgement was correlated with GM density in two areas (p<0.0001, uncorrected): one in the left IFG, and the second, in the post-central gyrus, extending from its inferior part (BA43) superiorly (figure 5.5). Age and time since stroke, when added as covariates, did not influence the results, while lesion volume, as before, resulted in no significant results. The analysis of the WM did not show any significant results when looking at inner speech alone.

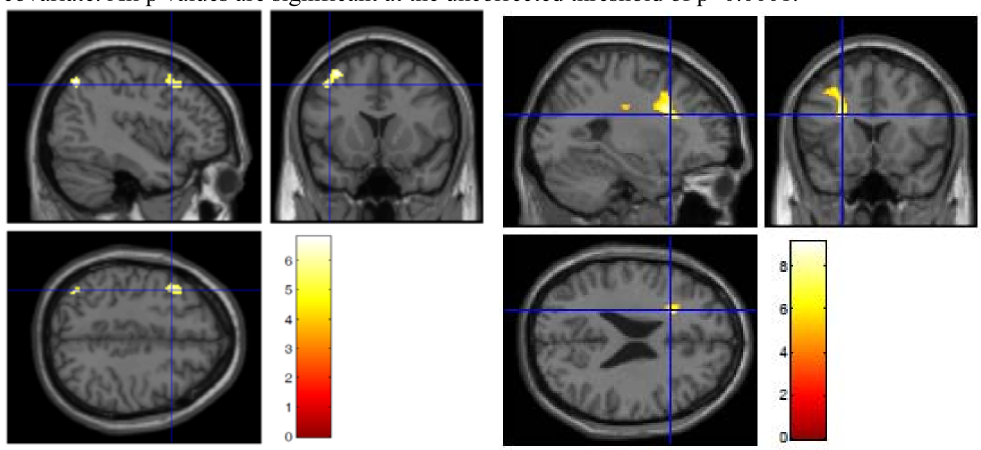
Figure 5.5. VBM maps showing z values of all grey matter voxels for the rhyme judgement, in the patient group. All p values are significant at the uncorrected threshold of p<0.0001.



One of the main differences between the homophone judgement and the rhyme judgement is that performance of the rhyme judgement task requires the use of working memory, while performance of the homophone judgement does not. To evaluate the contribution of working memory processes to the results, a measurement of verbal working memory (sentence repetition) was included as a covariate. After including this covariate no cluster survived the specified threshold (p<0.0001, uncorrected). The same results were obtained when adding word repetition scores as a covariate in the model.

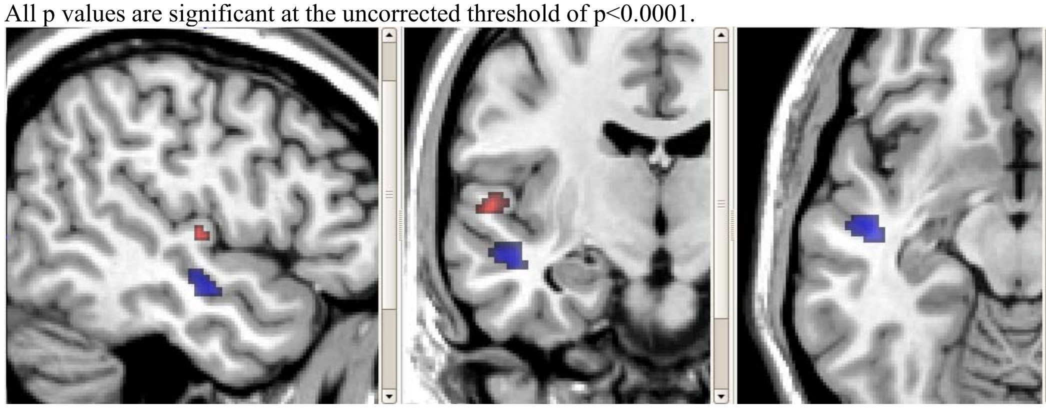
Subsequently, inner speech was compared to overt speech. Inner speech was significantly correlated with GM clusters in the left MFG, DLPFC regions (BA8/9/46) and the angular gyrus (BA39) (p<0.0001, uncorrected) over and above overt speech. Analysing the WM images, it was found that inner speech was significantly correlated with a white matter region extending from an area bordering the prefrontal (BA9/46) and premotor (BA6) regions, inferiorly towards the left IFG (BA44) (p<0.0001, uncorrected, figure 5.6).

Figure 5.6. VBM maps showing z values of all grey matter (left) and white matter (right) voxels for the rhyme judgement, for the patient group, when a measurement of overt speech was included as a covariate. All p values are significant at the uncorrected threshold of $p<0.0001$.



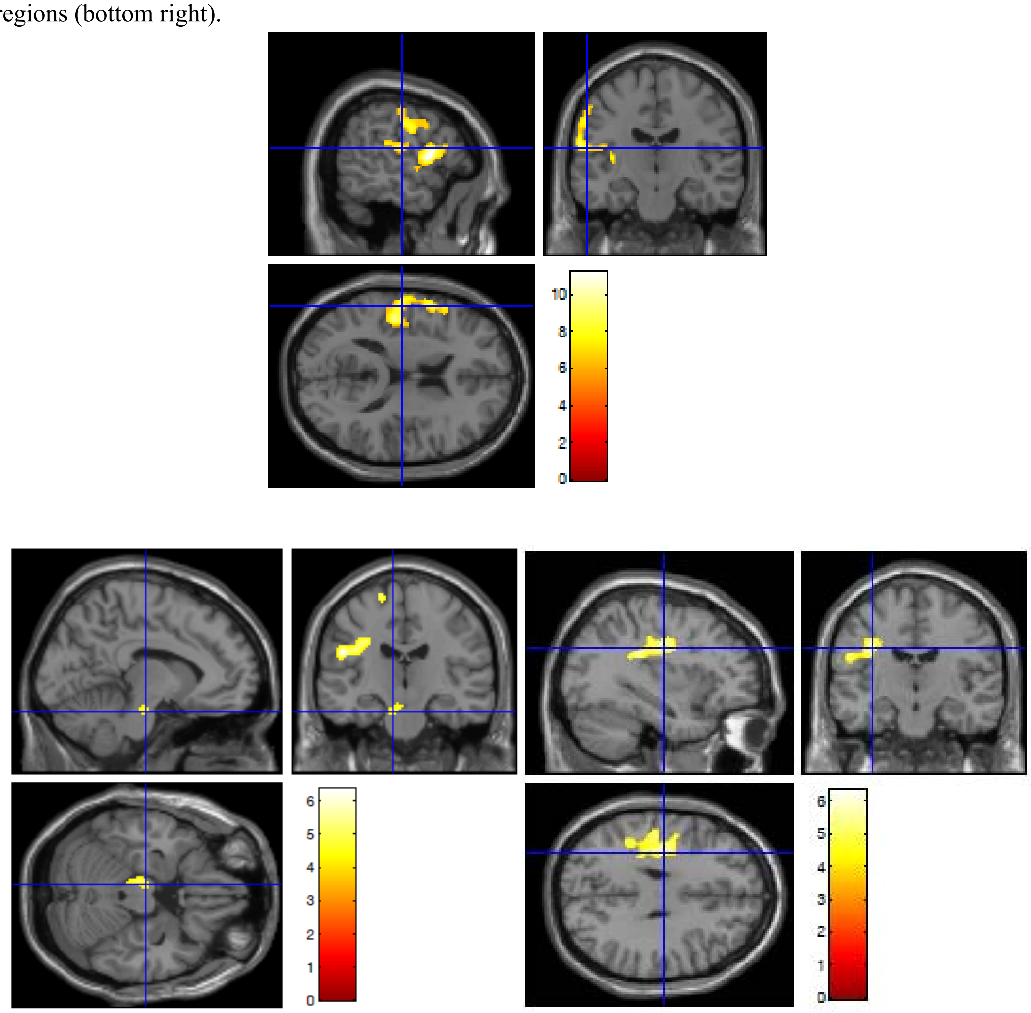
Looking at the entire group of controls (n=31), performance on the rhyme judgement was significantly correlated with GM density in the middle section of the superior temporal gyrus (BA22), and with WM density in a region inferior to it, in the superior temporal sulcus (STS) ($p<0.0001$, uncorrected). Interestingly, same results were obtained when including overt speech as a covariate (figure 5.7).

Figure 5.7. VBM maps showing z values of all grey matter (red) and white matter (blue) voxels for the rhyme judgement, for all controls, when a measurement of overt speech was included as a covariate. All p values are significant at the uncorrected threshold of $p<0.0001$.



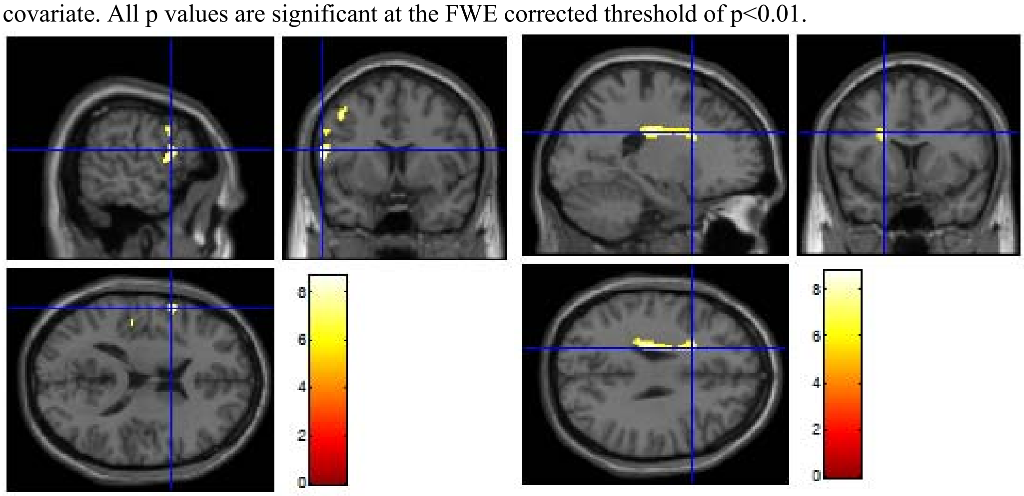
The scans of all participants (patients and controls, n=48) were then analysed, resulting in somewhat similar results to the ones obtained from the homophone judgement task. Namely, performance on the rhyme judgement task was significantly correlated with clusters in the left insula, IFG and post central gyrus (FWE correction, $p<0.01$). Few smaller clusters around the ventricles and the aqueduct were found as well, and as before, these disappeared when a group covariate was included (FWE correction, p<0.01). Next, the unmodulated white matter images of all participants were analysed. Interestingly, the results showed a very different pattern from those obtained for the homophone judgement. It was found that performance on the rhyme judgement correlated with WM density in three main areas (FWE correction, p<0.01): an area medial to the MTG (similar to that observed for the homophone judgement), the corticospinal tract (in proximity to the pons, left to the midline), and a frontal area medial and superior to the IFG, extending superiorly towards the middle frontal gyrus. When including a group (patients *versus* controls) variable, the two latter clusters, the corticospinal tract and the frontal area, remained significant (FWE correction, p<0.05). See figure 5.8.

Figure 5.8. VBM maps showing z values of all grey matter (top) and white matter (bottom) voxels for the rhyme judgement, for all participants (group covariate included). All p values are significant at the threshold of p<0.01 (FWE), for the grey matter images, and p<0.0001 (uncorrected) for the white matter images. White matter images show the corticospinal tract (bottom left) and frontal white matter regions (bottom right).



In subsequent analysis, the scores of reading aloud of the words used in the rhyme judgement task were included as a measurement of overt speech. It was found that GM density significantly correlated with inner speech scores in the regions of the IFG (frontal operculum and pars opercularis; BA44), and the DLPFC (BA9), over and above overt speech (FWE correction, p<0.01). Significant WM clusters were found in a region parallel to the left lateral ventricle, in the inferior parietal and frontal lobes (FWE correction, p<0.01). Because of the proximity of this region to the ventricles, it was further examined whether the result will be affected by adding a group covariate to the analysis. It was found that adding a group covariate did not change the result in this case.

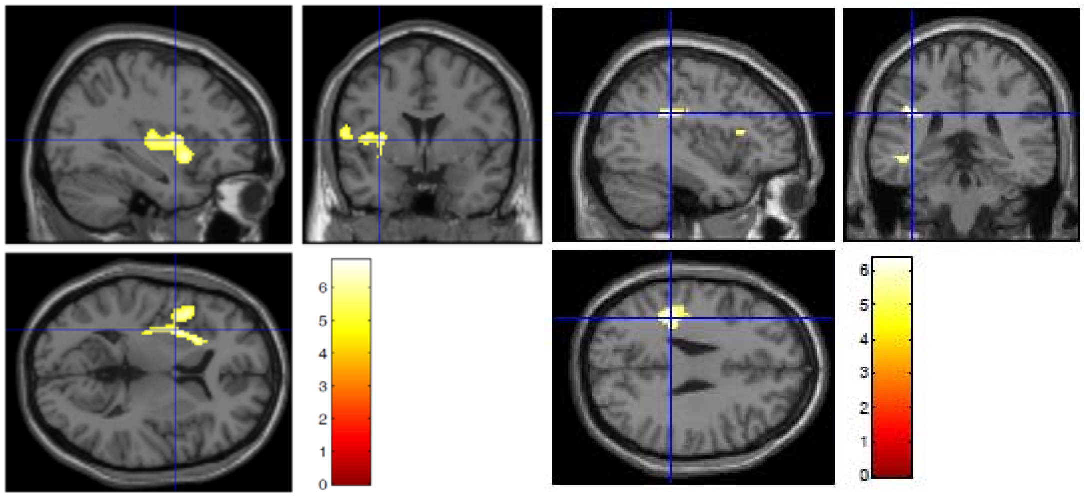
Figure 5.9. VBM maps showing z values of all grey matter (left) and white matter (right) voxels for the rhyme judgement, for all participants, when a measurement of overt speech was included as a covariate. All p values are significant at the FWE corrected threshold of p<0.01.



Lastly, the neural correlates of the **non-word homophone judgement** task were explored. Only 14 patients completed this task, and indeed, when analysing the patient data, no significant clusters were found. Similarly to the analysis of the homophone judgement, no significant results were obtained when analysing only the control data.

Analysing the data from all participants (n=45), it was found that performance on the non-word homophone judgement task was correlated (p<0.01, FWE) with grey matter density in the insula and IFG pars opercularis (BA44). Same results were obtained when adding a group covariate, just at the lower uncorrected threshold of p<0.00001. White matter clusters which significantly correlated with performance (p<0.01, FWE), were found in the inferior parietal lobule (IPL) and IFG. Similar results - although this time they also included a small cluster in the MTG - were found when adding a group covariate (p<0.00001, uncorrected, figure 5.10).

Figure 5.10. VBM maps showing z values of all grey matter (left) and white matter (right) voxels for the non-word homophone judgement, for all participants (group covariate included). All p values are significant at the uncorrected threshold of p<0.00001.



When adding a measurement of overt speech (reading aloud of the non-words), a small right cerebellar GM cluster and a left insular WM cluster were found to be significantly correlated with inner speech, over and above overt speech. However, both clusters were very small (27 and 25 voxels respectively).

3.3 The Effects of Ageing on Global and Language Related, Structural Changes

In this section effects of ageing were explored by analysing data from the healthy volunteers only. **Global effects of ageing** were examined by comparing the young (n=12) and older (n=19) controls (unpaired t-test, unequal variance assumed). Higher GM density in the younger controls compared to the older controls, was found in a few regions of the brain, including bilateral cerebellum, right pre-central sulcus, and the superior part of the left middle frontal gyrus (frontal eye field). Higher WM density was found in the internal capsule and the corticospinal tract, at the level of the pons (FWE correction, p<0.01).

The effects of ageing on performance on the **homophone and rhyme judgement** were examined for both the WM and GM images, both with and without the inclusion of overt speech as a covariate. No significant results were obtained in any of these analyses at the predetermined threshold.

3.4 Effects of Scanner Type

Adding a covariate which describes the scanner type did not affect the results of all analyses, excluding the one examining written sentence comprehension. For the written word comprehension, similar results were obtained, only at a lower threshold. Running the analyses with all participants, not including the three patients who were scanned on a different scanner, had a different effect on the different analyses. In the analyses of written and auditory comprehension of words and sentences (overall 4 analyses), no significant results were obtained when including only 17 patients. Analyses of the phonological fluency and reading aloud tasks resulted in similar results to the one obtained when including 20 patients. Lastly, the analyses of the entire population (homophone and rhyme judgement) were not affected by the exclusion of the three patients.

4. Discussion

This study looked at the neural correlates of inner speech, in healthy volunteers and stroke patients, using VBM.

4.1 Neural Correlates of Language Production and Comprehension

Looking at the various language tasks, some previous results were replicated. The **reading aloud** task was mainly correlated with a cluster in the STG and supramarginal gyrus, areas which constitute part of a system responsible for translating orthography to phonology (reviewed in Pugh, Mencl et al. 2000; Schlaggar and McCandliss 2007). Similar to the results of the VLSM analysis, **comprehension of spoken sentences** was correlated with a large cluster in the MTG, a result which replicates previous findings (for example, Bates, Wilson et al. 2003; Dronkers, Wilkins et al. 2004; Baldo and Dronkers 2007). **Comprehension of spoken words** was correlated with GM density in the left post-central gyrus (BA3), an area which is not usually implicated in speech comprehension studies. However, it should be noted that this cluster was small, comprising of only 38 voxels with an uncorrected threshold. In the **reading comprehension tasks**, the clusters in the left ITG and MTG were found to be significant. This result is consistent with the results of the previous chapter and also with other studies suggesting that these regions are part of the ventral stream of reading (reviewed in Pugh, Mencl et al. 2000; Schlaggar and McCandliss 2007). However, both reading comprehension tasks were also correlated with GM density in regions of the medial frontal lobe. The only frontal region previously reported to be correlated with various language comprehension tasks, is the dorsolateral prefrontal cortex (Bates, Wilson et al. 2003; DLPFC; Dronkers, Wilkins et al. 2004). However, the DLPFC lies slightly laterally to the areas found in this study.

It should be noted that due to the small number of patients, the results reported in this section were obtained using an uncorrected threshold. However, the pattern of results verifies that even with this uncorrected threshold, most results are meaningful and consistent.

4.2 Neural Correlates of Inner Speech

The analysis of the inner speech tasks revealed a functional network of three broad anatomical areas which support inner speech, over and above overt speech production. These areas were found when examining the patient population alone, control participants alone, and all participants together, and in the three inner speech tasks: homophone judgement of words and non-words, and rhyme judgement.

The first anatomical area is a perisylvian one which includes areas which are traditionally perceived as language areas. This includes BA44 in the IFG, and posterior regions located in the inferior parietal lobe and include the supramarginal and angular gyri. White matter clusters known to connect these two regions were found as well. These results largely replicate the ones found in the VLSM analysis, presented in the previous chapter. VBM showed that the grey matter in the STG and white matter tracts leading from the STG towards the inferior parietal lobe are involved in inner speech production as well. The second area involved in inner speech was found to be the DLPFC (BA9/46), an area traditionally associated with executive and other higher cognitive functions. A third area, BA6 which is a motor region overlapping with the premotor cortex in the MFG, was found as well. The potential contribution of each of these anatomically defined regions to inner speech production will be now discussed.

4.2.1 The Perisylvian System's Role in Inner Speech

The white mater areas which were found to be associated with inner speech production, over and above overt speech production, are proposed to be part of the dorsal route for language. White matter tracts originating in the temporal lobe, can potentially be a part of either the dorsal or the ventral route for language, since medial to the temporal lobe the ventral and dorsal routes for language partially overlap. Without a specialised technique (such as DTI) it is difficult to define precisely the functional importance of this WM area. However, further white matter regions were found in the region between the temporal and inferior parietal lobe, suggesting that these white matter tracts belong to the dorsal language route. Hence, these findings expand on those reported in the previous chapter. While both techniques (VLSM and VBM) showed the importance of inferior parietal regions, such as the supramarginal and angular gyrus, and that of the IFG, the VBM analysis expanded on the VLSM result by showing that other regions within the dorsal language route are related to performance on the inner speech task, including white matter regions and the STG, hence defining a perisylvian network for inner speech.

The role of the IFG (pars opercularis) and of the supramarginal gyrus were discussed in the previous chapter. Perisylvian and other language areas, which were found in this analysis and not found before, include the insula, angular gyrus and the STG/STS. In the literature, the anterior insula is usually mentioned with regard to two discrete, though not necessarily contradictory, roles. The first is phonological processing (Price and Friston 1997; Price 2000; Price, Devlin et al. 2005; Vigneau, Beaucousin et al. 2006; Shafto, Burke et al. 2007), and the second is articulatory planning (Dronkers 1996; Wise, Greene et al. 1999; Price 2000; although see also Hillis, Work et al. 2004). Some consider the insula to be part of the dorsal route for language (Hickok and Poeppel 2007). Therefore, its involvement in inner speech can be interpreted as either a phonological one (similar to BA44), or with a role more closely related to that of premotor areas, discussed below. The angular gyrus is considered to be part of the temporo-parietal system responsible for rule-based translation of orthography to phonology, during reading (reviewed in Pugh, Mencl et al. 2000; Schlaggar and McCandliss 2007). However, studies show that it is more likely to be activated when reading sentences than when reading words, presumably because of its role in semantics (Price 2000). Consistent with this hypothesis, VLSM studies have shown that lesions in the left angular gyrus are correlated with deficits in auditory sentence comprehension (Dronkers, Wilkins et al. 2004) and with reduced semantic variety in spontaneous speech (Borovsky, Saygin et al. 2007). In inner speech, the angular gyrus can support access to the correct lexical item and its subsequent phonological analysis, by providing the semantic component of the word (this could create the lexical effect, discussed in previous chapters, and potentially create a frequency effect as well). This hypothesis is supported by the fact that the analysis of the non-word task did not reveal any angular involvement. Lastly, the STG cluster is part of the primary and association auditory cortices. These regions are involved in phonetic discrimination (reviewed in Price 2000; Boatman 2004; Demonet, Thierry et al. 2005) and are the starting point of the dorsal language route when performing tasks like speech comprehension and repetition. Wise et al. (2001) have shown that the area anterior to Heschel's Gyrus in the left STS, an area with almost identical coordinates to those of the region found in this study, responded to hearing words but not to hearing sounds, therefore showing a speech-specific response. In the inner speech task, phonetic analysis is the end point of the process, followed only by making an overt decision (indicating whether the pair of words rhyme or not).

It is therefore suggested that the dorsal route for language is involved in the basic linguistic processes involved in inner speech production. First, phonological retrieval and word production is supported by the anterior perisylvian regions. Then information is transferred to regions which connect speech production and speech comprehension, such as the supramarginal gyrus. This information transfer occurs through white matter tracts which connect inferior frontal and inferior parietal/superior posterior temporal regions, and are most likely a part of the arcuate fasciculus. The angular gyrus meanwhile provides semantic support to both processes, and finally, phonetic analysis takes place in auditory association cortices in the superior temporal lobe.

4.2.2 The Inner Speech Executive System

The second region which was found to be important for inner speech processing in this study is the left DLPFC, located in the MFG. This was significant even when including overt speech production as a covariate in the analysis, and was replicated in analyses of both GM and WM. It should be noted that the MFG cluster in the different analysis largely overlap, as can be seen in figure 5.11. (Figure 5.11

shows clusters that were presented in the result section. It is added to ease visualisation).

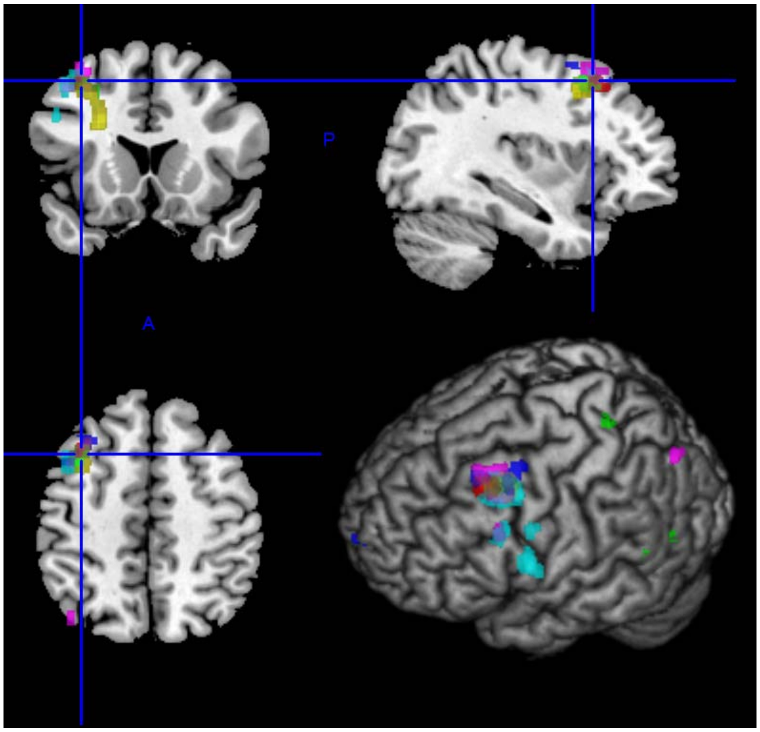


Figure 5.11.
Significant clusters in the MFG which correlate with performance on the inner speech tasks, over and above overt speech production. Clusters shown were found when analysing the homophone judgement and the WM images of the patient population (red), GM images of all participants (blue) and WM images of all participants (green), and when analysing the rhyme judgement and the GM images of the patient population (pink), WM images of the patient population (yellow), and GM images of all participants (turquoise).

Activation in the left DLPFC during inner speech production was previously shown in studies looking at verb and noun generation (Wise, Scott et al. 2001) and rhyme judgement (Pugh, Shaywitz et al. 1996; Lurito, Kareken et al. 2000; Hoeft, Meyler et al. 2007). Although these studies used inner speech tasks, they do not shed light on the specific role the DLPFC might have in inner speech production. The DLPFC is associated with the central executive of the working memory system (see a meta-analysis by Owen, McMillan et al. 2005), and with top-down modulation of attention or cognitive control (Kerns, Cohen et al. 2004; Egner and Hirsch 2005; Mansouri, Tanaka et al. 2009). With regard to the latter, the DLPFC is thought to divert attention and cognitive control specifically in cases of conflict, and it does so together with the anterior cingulate gyrus (ACC). Studies suggest that the ACC is in charge of recognising that two possible responses are inconsistent with each other or even contradicting each other. The ACC then sends a signal to the DLPFC which in turn diverts attention and other cognitive resources to the preferable response (Mansouri, Tanaka et al. 2009). The role of the DLPFC has also been implicated with regard to linguistic processes, such as lexical retrieval (Cappa, Sandrini et al. 2002;

Cotelli, Manenti et al. 2006; Fertonani, Rosini et al. 2009) and mapping of orthography to phonology and semantics (Liu, Hue et al. 2006). Clearly, some of these functions are involved in the performance of the homophone and rhyme judgement tasks. While reading a word pair such as BEAR and CHAIR, the orthography suggests that the words do not rhyme, since the two words end with a combination of different letters. However, the phonology shows that the words rhyme. These two responses are contradictory and one needs to be chosen over the other. This process might evoke the need for top-down modulation of attention which can assist in the resolution of the conflict. In addition, reading requires mapping of orthography to phonology, and all tasks in this study involved written material. It is therefore difficult to determine whether the involvement of this MFG region is related to one function over the others, or whether it is equally related to both functions.

4.2.3 The Motor System in Inner Speech

Involvement of the premotor cortex arises from this study. The MFG cluster described above bordered BA6 in some of the analysis. This might be a result of the compromised resolution and spatial sensitivity of VBM, which happens especially when studying stroke patients and when applying spatial smoothing, as will be further discussed below. However, the potential involvement of this brain region is considered as well. Another area, related to motor planning repeatedly found in this study is the left insula. Left premotor activation was found in studies which examined inner speech production using a noun and verb generation task (Wise, Scott et al. 2001), rhyme judgement (Paulesu, Frith et al. 1993; Owen, Borowsky et al. 2004), semantic fluency (Basho, Palmer et al. 2007) and silent counting (Ryding, Bradvik et al. 1996). However, it should be noted that BA6 covers large area of the MFG, and in some of these studies the exact coordinates in which peak activation was found are somewhat distant from the peak coordinates found here. Lastly, some scholars consider the premotor cortex to be part of the dorsal route for language (Wilson, Saygin et al. 2004; Hickok and Poeppel 2007; Saur, Kreher et al. 2008). Its specific roles should therefore be explored.

Both the premotor cortex and the insula are involved in motor planning. The premotor cortex is active during both execution of movement and during motor imagery in stroke patients (Sharma, Simmons et al. 2009) and in healthy participants (Solodkin, Hlustik et al. 2004; Sharma, Simmons et al. 2009), while the insula is involved in planning articulation (Dronkers 1996; Wise, Greene et al. 1999). It is suggested that the role of these motor regions in inner speech production might be related to the role of motor feedback in control of behaviour.

One of the main differences between inner and overt speech is in the feedback available to the speaker. This issue was discussed before and will be briefly reviewed again. Current models of speech production and reading aloud emphasise the importance of feedback (Postma 2000). Research suggests that after a linguistic structure (the desired word or sentence) is generated, three types of feedback become available: 1.) inner speech or the articulatory buffer (Levelt 1983; Levelt 1989; Levelt 1993), 2.) feedback from motor articulation (MacNeilage 1970; Kelso 1982; Levelt 1989; Guenther 2005), and 3.) feedback from overt speech via the auditory comprehension system (for a review see Postma 2000; Guenther 2005). While overt speech clearly 'enjoys' all three types of feedback, inner speech might only benefit from the first type of feedback, which in turn, might compromise performance to a large extent. Therefore, being able to make use of other sources of feedback which are available during overt speech, can potentially improve performance. Inner speech is defined as speech for which external auditory feedback is not available. Therefore, the only feedback stream which can potentially benefit inner speech is motor feedback. Since inner speech is not necessarily accompanied by actual movements of the articulatory apparatus, motor feedback can be achieved by creating motor commands which correspond to the word produced by inner speech. These motor commands can then be the source of feed forward and feedback mechanisms, even in the complete absence of executed movements.

Guenther (2005) suggests that areas involved in feedback hold *error maps* - pre-specified definitions of the correct sound and somatosensory input corresponding to the production of a specific word or syllable. These areas receive feed forward information when speech is produced, compare the output to these *error maps*, and send feedback to the speech production areas. For well-learned syllables, error maps can be also compared with the motor plan, not only with the actual motor execution. The premotor cortex holds plans of motor commands related to specific syllables, plans which are also known as *gestural score* (Browman and Goldstein 1992) or *mental syllabary* (Levelt and Wheeldon 1994). These motor plans can be used in a feed forward mechanism. In this process, the stored plans are compared to the actual planned movements, and actual motor commands are corrected when needed

(Guenther 2005). This idea is supported by Kelso (1982), who agrees that motor plans are monitored even before execution, through a 'central efference monitor'. In a study which looked at inner speech monitoring, participants were asked to produce 'tongue twisters' and report the number of self-corrections (Postma and Noordanus 1996). Participants repeated the task in different conditions: inner speech, mouthing, overt speech in the presence of white noise and overt speech without noise. Interestingly, there was no difference in the number of errors detected by the participant in the first three conditions. These results support the idea that the feed forward mechanisms are related to motor planning (rather than execution) and that speech sounds can be active even in the absence of articulation or auditory feedback.

It is suggested, however, that motor feedback loops are not essential for the production of inner speech, but that they can boost performance. It might even be the case that their use is task-dependent: subjects can choose to use this extra source of feedback when monitoring of overt speech is important, and not to use it when inner speech can be compromised. For example, when thinking to oneself, the semantic, phonological, and even grammatical accuracy of inner speech is not essential. In this case, motor involvement might not be needed. However, when performing a task like the homophone and rhyme judgements used in this study, production accuracy is essential for task performance. In this case, motor feedback might greatly contribute to performance. This idea is supported by the results of a study which was already mentioned before, in which participants vocalised more as the difficulty of reading material increased, and had lower comprehension scores when prevented from vocalising or even sub-vocalising (Hardyck and Petrinov 1970). McGuire et al. (1996) manipulated inner speech automaticity by giving participants two tasks: In the first they had to say simple sentences in their head, while in the second task, which was rated by participants as harder, they had to imagine someone else saying the same sentences. The first task is therefore more automatic and cognitively less demanding. Comparing the two conditions directly revealed that the more difficult task resulted in greater activation in bilateral SMA and premotor regions. The authors' explanation to this finding is that when inner speech requires more accurate output, motor regions are involved more, and they further suggest that the generation of these motor plans is closely linked to the generation of phonological forms (McGuire, Silbersweig et al. 1996). It should be interesting, therefore, to measure premotor involvement (with fMRI or PET) while manipulating the importance the subjects give to monitoring their own inner speech. This can be done, for example, by giving three different groups of subjects to read silently the same poem, and scanning them while they read. One of the groups would be told that they will be asked about the structure, rhyming, or other relevant aspects of the poem, later, and that they will be paid for every correct response. The second group would be told that they will be asked about the text, and the third group would not be given any specific information. If successful, this manipulation will affect the motivation of the subjects in the different groups, therefore causing the first group to put the most emphasis on monitoring their inner speech, and the third group, the least. Accordingly, one can hypothesise that performance and premotor activation will vary between the groups, with the most motivated group showing better performance and greater premotor activation.

An alternative explanation for the involvement of the premotor cortex comes from a study by Solodkin et al. (2004) which suggested that the role of the premotor cortex during motor imagery is to suppress the execution of movements. Basho et al. (2007) compared inner and overt speech during a phonological fluency task. The authors have suggested that activation in the left superior frontal gyrus (BA6), right cingulated gyrus (BA32), right superior frontal gyrus (BA11) and right inferior and superior parietal lobe (BA40 and 7), might be attributed to inhibition of overt response and response conflict (producing a word but not saying it aloud). However, this explanation is relatively unspecific, and was not further explored in these studies or following studies. In the VBM analysis, grey matter density in the premotor cortex was correlated with behaviour. If the role of the premotor cortex is solely to inhibit response, then the interpretation of the results should be that greater inhibition correlates with better performance. However, there is no theoretical reason to support this interpretation. Moreover, a previous study presented above, has shown the exact opposite effect: better performance was correlated with less motor inhibition (Hardyck and Petrinov 1970). In summary, the premotor cortex might or might not be involved in motor inhibition during imaginary movement. But whatever the case is, motor inhibition cannot explain why grey matter density in the premotor cortex was correlated with performance in this study.

Another motor region which was found to significantly correlate with performance across all participants was the white matter of the cortico-spinal tract. This region correlated with performance on the rhyme judgement task when looking at inner speech but not when adding overt speech as covariate. The cotrico-spinal tract is involved in sending signals from the cortex to the articulatory muscles. This result reflects the fact that the cortico-spinal tract is involved in execution of motor commands, rather than motor planning.

Lastly, one might suggest that the correlation between tissue density in motor regions and performance on the inner speech task is an epiphenomenon, resulting from the fact that patients with severe language deficits are in general those with bigger lesions, which inevitably encompass motor regions as well. Indeed, performance on all three inner speech tasks was significantly correlated with levels of speech apraxia (Kendall's tau =0.43-0.49, $p<0.005$, for all three task) as determined using the Apraxia Battery for Adults. If motor regions are significant solely due to the correlation between inner speech performance and motor abilities, one would expect to see involvement of motor regions also in other language tasks on which performance is correlated with speech apraxia. Nonetheless, similar or higher correlations were obtained between speech apraxia scores and scores of auditory sentence comprehension and reading aloud words, for example, but these tasks were not correlated with tissue density in motor regions. It is therefore difficult to argue that the clusters which were found to be significant in motor regions are merely a secondary outcome.

4.2.4 The Influence of Working Memory on Inner Speech

When exploring the influence of working memory, as measured by the sentence repetition task, on the neural correlates of rhyme judgement, no significant results were found. When the influence of word repetition was examined, no significant results were found as well. This is consistent with the findings from the VLSM study presented in the previous chapter. It was argued before that word repetition has substantially lower working memory load compared with sentence repetition. Accordingly, one would expect the effect of word repetition and sentence repetition to be different. That is, if including sentence repetition as a covariate should remove the working memory areas from the statistical map which represents the inner speech areas, including word repetition as a covariate should not have this effect. However, this was not the case, and as mentioned before, word and sentence repetition had the same effect on the data. It is therefore suggested that the lack of significant results is not because working memory explains all of the variability in the rhyme judgement task. Rather, it might be because both inner speech and repetition rely on the dorsal language stream. It is difficult to fully disentangle repetition from working memory, especially because in clinical settings, working memory tests traditionally employ repetition (e.g. digit span, sentence repetition). In order to fully understand the contribution of working memory to the results obtained from the rhyme judgement task, one should employ a working memory task which does not involve repetition.

4.3 Methodological Aspects of the Current Study

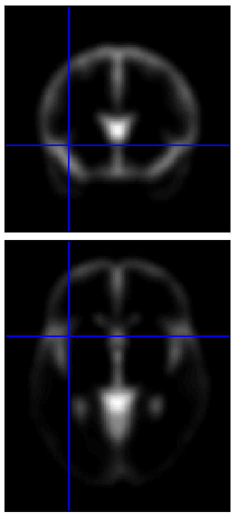
Some of the analyses presented in this chapter did not yield significant results. In the case of the homophone judgement (both words and non-words) of the control participants alone, and of the ageing effects in the context of language performance, the lack of significant results might be attributed to the very small variability in the behavioural data in the control sample, which reflects a ceiling effect. When there is very small variance within the behavioural data it is very difficult to find structural differences. This is clearly not the case with functional data. Many studies show that equal performance might be supported by different functional mechanisms. Moreover, today many scholars in the area of functional imaging, argue that it is of no use to even attempt at comparing brain activations between two populations if their behavioural performance is not equal (for example, see Price and Friston 1999). The analyses of the non-word homophone judgement did not yield any significant results in two cases as well: when analysing the data from the patients only, and when analysing the data from all participants and including a covariates of overt speech production. In both cases, this might be due to the small number of patients who completed the non-word homophone judgement task, a number which can potentially substantially reduce the power of the analysis.

Another problem in this study was the fact that three patients were scanned on a 1.5T scanner while the rest of the participants were scanned using a 3T scanner. This can potentially introduce noise, or even systematic errors, into the data. The source of such noise can be, for example, the difference in the magnet field strength, noise of the electronics of the MRI which can vary between scanners, and differences in subjects' positioning in the scanner (Stonnington, Tan et al. 2008). Here it was impossible to systematically evaluate the influence of these factors on the data. Moreover, a comparison between scanners could not be achieved as well because of the small number of subjects scanned on one of the scanners. To obtain some estimation of the influence of this variable, two analyses were adopted. In the first, scanner type was added as a covariate to the general linear model, and in the second, the three subjects scanned on a different scanner were removed. The results were then qualitatively compared to the full dataset. This is by no means the ideal way to estimate the effects of scanner type on the results of a multisite study. However, given the available data and the fact that this study was not a full multi-centre study, this was a feasible method. The results show that the influence of scanner type, if exists, is negligible. This conclusion is supported by a previous study showing that volumetric measurements did not differ between a 1.5T and a 3T scanner (Briellmann, Syngeniotis et al. 2001). Moreover, Stonnington et al. (2008) have shown that in a multi-centre study the effect of disease (in this case, Alzheimer's disease) was far greater than the effect of scanner, and they go further to suggest that pooling data from 1.5T and 3T scanners should be viable in VBM studies which use SPM5 and above.

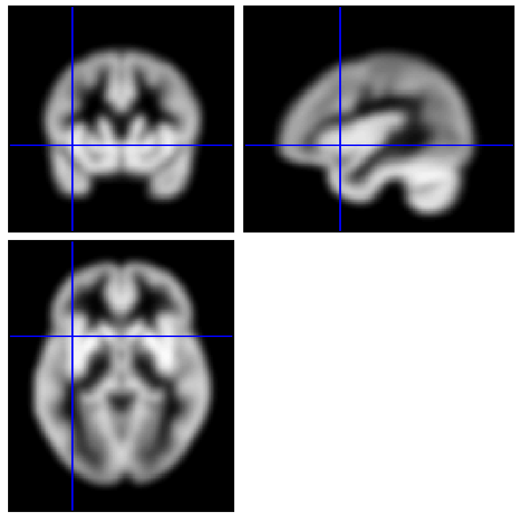
Although VBM has many advantages - probably, the most important being that it is an automated technique which covers the entire brain - it also has some disadvantages. It has been mentioned above, that recent advances in normalisation and segmentation procedures allow analysing lesioned brains. However, normalisation of abnormal brains, although improved in SPM5 onwards, can still result in some local distortions (Mechelli, Price et al. 2005). For example, if subjects have abnormally large or distorted ventricles (as is the case with many stroke patients), differences in structures on the borders of the ventricles might be an artefact of the normalization procedure, since structures close to the ventricles might be displaced during normalization (Mechelli, Price et al. 2005). This was indeed the case in this study. It was found that clusters along the ventricles usually disappeared, or at least were substantially minimised, when adding a group covariate to the analysis.

Another source of distortion in the data can be found by a close examination of the segmented images, which revealed that in cases where the lesion is large and old, the segmentation procedure is not perfect. Specifically, the core of the lesion, which often develops into a liquid filled cavity, might be classified as CSF, although it is an area which before the stroke 'held' grey or white matter. As a result, parts of the lesion are actually not included in the GM/WM analysis, and their importance for the language function in question is overlooked. In VLSM, on the other hand, the core of the lesion is clearly abnormal and will almost always be classified as a lesion. This can partly explain the discrepancy in results between the VLSM and the VBM analyses. Figure 5.12 demonstrates this segmentation problem. In the figure, the *a priori* templates of CSF, GM and WM can be seen, and are compared to a GM image of one patient. The values in each voxel represent the chance of this voxel belonging to a specific type of tissue. It can be seen that in this image, a voxel which is in the core of the lesion, is classified as more similar to CSF than to GM.

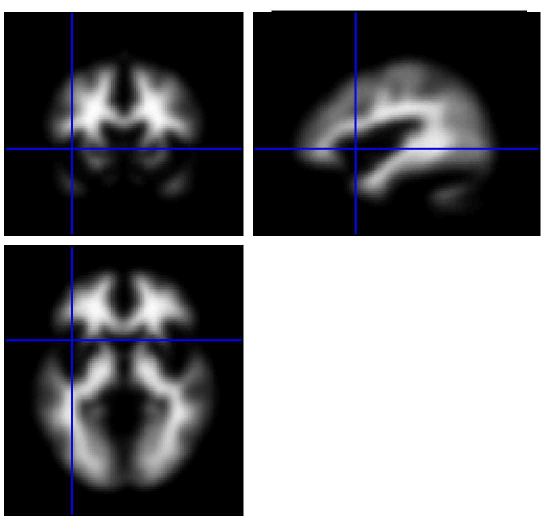
Figure 5.12. *A priori* images and a segmented GM image of one patient. The values of the voxel in the various images are: CSF=0.116, GM=0.845, WM=0.029, patient=0.092.



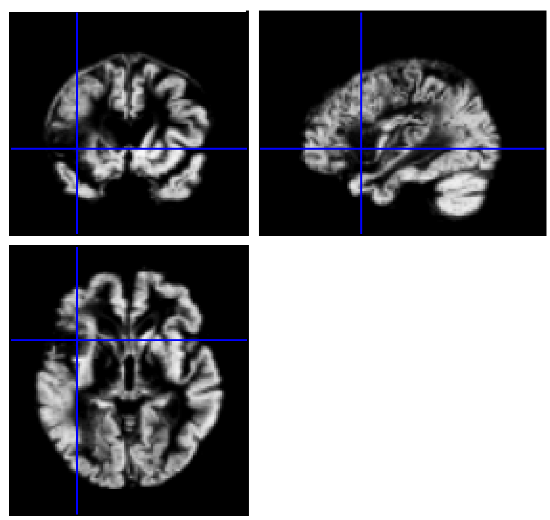
A priori CSF template



A priori GM template



A priori WM template



Segmented GM image of a patient

The next stage in the pre-processing of the image, smoothing, creates difficulties for the analysis as well. Apart from the obvious issue of reducing spatial resolution, a problem which is not special to VBM, smoothing also shifts the local maxima slightly towards areas of low tissue variance, thus further distorting the map, and makes it difficult to give accurate localisation (Mechelli, Price et al. 2005). Lastly, there are problems with interpretation of voxel intensity. In lesioned brains, the relation between voxel intensity and the fate of the tissue is not fully understood. For example, it has not been examined whether higher intensity means a larger number of functioning neurons. Differences between stroke patients and healthy volunteers are more easily understood since they can be attributed to the large abnormalities in the patients' brains, reflecting tissue loss. With regard to normal healthy brains, it is unclear whether differences observed in healthy subjects are related to changes in neuropil, neuronal size, dendritic arborisation, axonal arborisation or other morphological differences.

4.4 Conclusions and Further Research

In this chapter, VBM was used to explore the neural correlates of inner speech. Data presented here corroborated the main findings of the VLSM study but has expanded on them to enable the construction of a more complete model of the inner speech network. Three main functional systems were found to correlate with inner speech performance. The first one is a perisylvian system covering regions involved in the dorsal route for language. In this system, inner speech is produced by frontal regions (BA44 and possibly the insula), transferred via the arcuate fasciculus to posterior regions which link speech production to speech comprehension and to semantics (the supramarginal gyrus and the angular gyrus, respectively), and finally speech output is analysed by the auditory association cortex. DTI studies are needed to further understand the connectivity patterns between these different brain regions.

The second system is one which has an executive function on inner speech production. The DLPFC is suggested to be responsible for conflict resolution. Previous studies have also suggested that it has a linguistic function, but this has not been reliably replicated in many studies. Lastly, a premotor system is suggested to have a feed forward role, the importance of which might be dependent on task demands, although this suggestion clearly deserves more investigation.

Except for the methodological limitations discussed above, VBM is also limited in its ability to correlate brain structure to cognitive function. Specifically, structural studies usually reveal those areas which are necessary for a specific cognitive function, rather than the ones which simply contribute to the behaviour. Functional imaging complements structural studies in that it can recognise the areas sufficient for a task. Moreover, in cases where compromised tissue integrity might initially cause a decline in performance, functional imaging can potentially uncover compensatory brain mechanisms. For example, in this study some analyses did not produce any significant results, and it was suggested that this is due to the small variability in behaviour. However, same external behaviour can be supported by very different brain mechanisms, a fact which cannot be revealed by structural studies, but can be revealed by functional studies. The next chapter, therefore, presents an fMRI study of inner speech, with the aim of exploring compensatory mechanism which are related to stroke and ageing.

Chapter 6
An fMRI Investigation of Inner Speech

"The brain itself is an excessively vascular organ, a sponge full of blood, in fact; and another of Mosso's inventions showed that when less blood went to the arms, more went to the head. The subject to be observed lay on a delicately balanced table which could tip downward either at the head or at the foot if the weight of either end were increased. The moment emotional or intellectual activity began in the subject, down went the balance at the head-end, in consequence of the redistribution of blood in his system.

William James, The Principles of Psychology (1890)

1. Introduction

PET and fMRI have been used to study language processing since the early 1990s'. Today, functional imaging of language is one of the more prolific areas in cognitive neuroscience research. While the VLSM and VBM studies, presented in the previous chapters, provided evidence of the areas necessary for inner speech, lesion studies by themselves cannot reveal the full extent of the areas sufficient for a given cognitive process. Moreover, structural studies cannot identify compensatory mechanisms which are used to restore behaviour after brain damage or during ageing. Lastly, impairment in inner speech can also be related to altered functional activity, which cannot be seen structurally with the available techniques and their resolution. For these, functional imaging studies are used.

1.1 Inner Speech in Functional Imaging Studies of Aphasia

Studies exploring inner speech and its relation to overt speech using functional imaging were reviewed in chapters 4 and 5. Previous studies which used rhyme judgement replicated the importance of the left IFG and the supramarginal gyrus in inner speech production and phonological analysis. Studies which compared inner and overt speech showed a distributed network of regions which included the left perisylvian and frontal regions, as well as right hemispheric language homologues and the right cerebellum. It was also noted that some of these studies suffer from a number of methodological caveats, the main one being the use of tasks which give no indication regarding the participant's behaviour and performance during the scan.

This study was designed to characterise the neural correlates of inner speech in healthy volunteers and aphasic patients. It differs from the last two chapters, in that this study looked at the function, rather than at the structure of the neural systems which supports inner speech. Two previous studies looked at the functional correlates of inner speech in post-stroke aphasia. In the first study (Calvert, Brammer et al. 2000), a 28-year old stroke patient performed a non-word rhyming task with nearly 100% accuracy. One large cluster of activation was found in the right inferior frontal cortex. Other areas, which were reliably activated in 3 healthy age-matched participants, including the left middle frontal gyrus, the anterior cingulate and the posterior border of the lesion in the region of the left IFG, showed reduced activation in the patient. The authors concluded that residual activation in the left IFG as well as recruitment of homologous areas in the right hemisphere, supported the patient's recovered performance (Calvert, Brammer et al. 2000). In the second study (Perani, Cappa et al. 2003), 4 patients with aphasia and 10 healthy participants (not age-matched) performed phonological and semantic fluency tasks. During the phonological fluency task, healthy participants showed left hemispheric activation of the IFG, DLPFC, the inferior parietal lobule, the basal ganglia and the thalamus. In the right hemisphere, the inferior frontal cortex and the thalamus were activated. Bilateral activation was found in the retrosplenial cortex and the pre-supplementary motor area (pre-SMA). During the semantic fluency task, healthy participants showed left hemispheric activation in the IFG; pars triangularis, DLPFC, MTG, occipitoparietal junction, retrosplenial cortex and cuneus. A patient with aphasia, who had complete recovery and a stroke not involving frontal areas, showed activation comparable to that of healthy controls. In a second patient, with a lesion involving part of the IFG, activation was similar to controls but also involved more right hemispheric activation. Lastly, patients who showed very little recovery, and who had extensive damage in the frontal and temporal areas, showed very little to no activation in the different tasks performed (Perani, Cappa et al. 2003). It should be noted that when the task was performed covertly during the fMRI scan, no behavioural measurements were recorded. It is therefore unclear whether the performance using inner speech was comparable to that using overt speech. Moreover, no data regarding apraxia were presented, and it is therefore difficult to determine to what extent the language production deficit is a result of motor dysfunction.

1.2 Age Related Changes in Brain Activation of the Language System

A discussion of the changes in brain structure and cognitive abilities, brought about by ageing, was presented in the previous chapter. Cognitive deterioration is not equal in all cognitive domains, with memory usually being affected the most by ageing. In the area of language, word retrieval is most affected by ageing while comprehension usually remains intact even in very old age. Ageing also does not affect all brain regions equally, and previous studies suggest that the frontal lobes show the largest age related reduction in grey matter density.

Researchers attempt to explain cognitive changes in terms of changes in brain structure and function. In the previous chapter a single VBM study was presented which showed that language production deficits correlated with grey matter atrophy in healthy ageing participants (Shafto, Burke et al. 2007). In the area of functional imaging, recent years have brought about a dramatic increase in the number of studies of ageing and cognition (Grady 2008), and a number of fMRI studies have explored the relation between speech production and ageing (Wierenga, Benjamin et al. 2008; Galdo-Alvarez, Lindin et al. 2009; Meinzer, Flaisch et al. 2009; Shafto, Stamatakis et al. 2010). The following section will describe these studies within the framework of some of the main theories in the area of cognition and ageing.

A current theory claims that the structural and functional integrity of the cortex in general, and of the frontal lobes in particular, decrease with age, causing the observed decline in performance (West 1996). Cabeza (2002), on the other hand, suggests that the reduction in hemispheric lateralisation seen in older adults is the mechanism underlying compensation for cognitive decline. This theory is summarised in the *Hemispheric Asymmetry Reduction in Older Adults* (HAROLD) model. Others suggest that decline in processing efficiency and compensation can be demonstrated in various areas of the brain, therefore not committing to a specific anatomical theory. In general, when normal levels of performance are accompanied by increased activity or activity in a novel area, it is usually interpreted as evidence for compensatory mechanisms. On the other hand, when impaired performance is accompanied by increased or normal activation, it is usually interpreted as processing inefficiency (reviewed in Grady 2008). However, other researchers suggest that wider activation can reflect the use of different cognitive strategies (Grady 2008). Wierenga et al. (2008) found that right hemispheric activation during picture naming was more extensive in older adults. Right IFG activation was positively correlated with performance, interpreted by the authors as a compensatory mechanism, and further supporting the two theories presented above (West 1996; Cabeza 2002). However, right pre-central gyrus activation was negatively correlated with performance, providing evidence for processing inefficiency. The study shows that both beneficial compensation and disadvantageous inefficiency can occur in older age, in the same subject and in the same task (Wierenga, Benjamin et al. 2008). In another study, participants performed a semantic and phonological fluency task. In the phonological task, behaviour and brain activations were similar for the two groups. In the semantic task, on the other hand, older adults had poorer performance and this was accompanied by wider activation in both hemispheres. Moreover, right IFG and MFG activations were negatively correlated with performance (Meinzer, Flaisch et al. 2009). These results are interpreted as demonstrating inefficiency as well. Lastly, following their VBM study (Shafto, Burke et al. 2007) which showed ageing effects related to word retrieval in the left insula, Shafto et al. (2010) examined fMRI activation during picture naming. They found left insula activation in trials where participants knew the answer but also when they incurred a TOT state. Moreover, the left insular activation was not affected by age during successful retrievals, but during TOT states, activation was lower for the older group, in comparison with younger controls.

Another prominent theory called '*cognitive slowing*' (Salthouse 2003), suggests that with age there is a general slowing down of cognitive processes. A recent ERP study has shown that older adults are slower in naming famous people, and that ERP signals associated with early visual processing and encoding did not differ between the two age groups. However, looking at later signals, it was shown that the older group had a delay in the signals associated with face recognition and semantic-lexical retrieval (Galdo-Alvarez, Lindin et al. 2009). The authors suggest that these findings support Salthouse's theory of 'cognitive slowing'. They emphasise that the 'slowing' is specific to some processing stages but not to others. The study also showed that young and older adults differ not only in the timing of their brain signal but also in its amplitude and distribution: older adults showed smaller amplitude in trials of successful naming and larger distribution of signal across the scalp. The latter finding is interpreted within the framework discussed above, which explains large signal distribution as evidence of a compensatory mechanism (Galdo-Alvarez, Lindin et al. 2009).

1.3 Current Research

Data regarding functional brain mechanisms underlying inner speech are scarce and incomplete. The existing data focus mainly on young healthy volunteers, and data regarding older subjects and aphasia patients hardly exist. The current study was aimed at characterising inner speech in these distinct populations, by using a rhyming task of both written words and pictures. The study employed a task requiring a response and incorporated a design that distinguished between the use of different cognitive strategies. The purpose of this was to overcome some of the caveats in previous studies. In order to minimise motor movement during inner speech condition, participants practiced the task outside the scanner until satisfactory performance was achieved. Lastly, in contrast to many previous studies which used a fixation cross as baseline, this study employed a high level processing task as a baseline. Previous meta-analysis has shown that this generates higher sensitivity to activation in relevant brain areas (Price, Devlin et al. 2005).

The research questions were applied to the three populations examined. First, the neural substrates underlying normal inner speech during word reading and picture naming were examined. The following questions were explored:

1. Does inner speech require both the anterior and posterior language areas? It was hypothesised that the inner speech tasks will generate widespread activation in the areas identified in the previous chapters, namely the left IFG, supramarginal gyrus, angular gyrus, insula, DLPFC and BA6. It was hypothesised that activation would also be found in areas responsible for executive function and motor control. Specifically, activation in the anterior cingulate cortex (ACC) might be seen during the word rhyming task, since the ACC works in concert with the DLPFC during conflict resolution. Activation was also expected in the SMA which together with the premotor region, is responsible for movement planning.
2. What are the areas common to inner speech following word reading and picture naming? This question can potentially differentiate areas which are essential for inner speech *per se* from areas which are more related to lead-in processes

(Indefrey and Levelt 2004). This question was addressed by using a conjunction analysis (Price and Friston 1997).

3. Based on the idea of motor feedback raised in the previous chapter, it was hypothesised that premotor activation will correlate with inner speech performance. In order to avoid the use of dependent data in the analysis a premotor ROI was defined around the local maxima found in the VBM analysis.

Looking specifically at ageing effects, the main research question was: What are the differences in brain mechanism underlying inner speech in younger and older healthy subjects? Based on previous studies, it was hypothesised that older adults will show widespread brain activation, in comparison to young adults, especially in right hemispheric frontal areas. If performance did not decline with age, then wider activation could be interpreted as a compensatory mechanism facilitating successful retrieval. If, on the other hand, older and younger adults differ in their behaviour, then wider activation might be attributed to inefficiency in processing. To distinguish inefficiency and compensation, correlations between these activations of interest and performance were examined. It was predicted that any differences between age groups might be more pronounced during the picture task, compared to the written word task. This is because difficulties in retrieval should influence task performance when pictures are involved, more than when written words are used.

Lastly, by examining the patient population, the study asked whether chronic patients show ipsi-lesional and/or contra-lesional activation. It was hypothesised that there would be a difference in brain activation between patients with cortical damage and those with sub-cortical damage. Patients with sub-cortical lesions were expected to show activation similar to that of the control group, while patients with cortical lesions, and especially those affecting perisylvian regions, were expected to show more widespread activation, potentially involving right hemispheric language homologues.

2. Methods and Materials

2.1 Participants

Six patients (Patients AD, BA, CB, DB, IB and LM, 4M/2F; age range: 49-78, mean age: 64.3±9.6; mean number of years of education: 14.3±3.7; mean time since stroke: 15.2±6.4 months) underwent an fMRI scan. Inclusion criteria, other than having no contra-indications to 3T MRI, included the ability to perform the inner speech tasks, being right handed, having a history of only one stroke, and not being on any medication other than that given for further stroke prevention. Appendix 2, table 1, gives demographic and other related information about the six patients who participated in the fMRI study. These patients are part of the cohort which participated in the studies of previous chapters. The control participants were the same as the ones who participated in the VBM study, presented in chapter 5. In short, two groups of controls participated in the study: 12 young controls (4M/8F; age range: 21-34, mean age: 24.6±4.5; mean number of years of education: 18.1±2.1) and 19 older controls (8M/11F; age range: 55-71, mean age: 64.1±4.8; mean number of years of education: 15.1±2.9). One young female revealed that she was ambidextrous after the scan and was therefore excluded from subsequent analysis. All healthy volunteers had no previous history of stroke and no history of neurological, psychiatric or language disorders. They were right handed and native speakers of British English. The study was approved by the Cambridge Research Ethics Committee and all participants read an information sheet and gave written consent.

2.2 Materials

To create stimuli for the rhyme judgement task, 130 nouns were chosen, which create pairs of rhymes according to the Oxford British Rhyming Dictionary (Upton and Upton 2004). The words' ending in each pair differed in their orthography, so that participants could not make the rhyme judgement based on orthography alone. For each noun, a black and white photo was created (360x362 pixels, white background). 30 native British-English speakers (14M/16F; age range: 23-63, mean age: 35.7±12.7; mean number of years of education: 18.1±3.8) were asked to name the pictures using one word. Pictures were presented in four blocks and the order of blocks was counterbalanced between participants. Pictures with high naming agreement (≥95%)

were chosen for the fMRI study. The final list contained 36 word pairs, out of which 26 pairs rhymed (about 70%) and the rest did not rhyme (about 30%)[13]. These words were presented as written words in one block, and as pictures in another. A potential methodological problem was that if each word appeared once in a rhyming pair and once in a non-rhyming pair, participants could develop a strategy whereby they might not use inner speech for some trials, but rather remember that if they saw the word before in a rhyming pair, it is necessarily a non-rhyming pair this time, and vice versa. To avoid this, some words appeared only in rhyming pairs (once or twice) and some words appeared once in a rhyming pair and once in a non-rhyming pair. This way participants could not predict the correct answer from the previous appearance of the word.

A high level baseline condition employing a visual similarity task, was chosen to control for activation associated with: a.) Visuo–spatial processing, b.) Comparison between two items, c.) Decision making, and c.) Motor response. Meaningless words and symbols were used and participants had to indicate whether two images were identical or not. The stimuli did not resemble letters or familiar pictures, therefore reducing the risk of inducing spontaneous naming. For the words baseline, 26 symbol strings were used. In 26 pairs the two symbols were identical, and in additional 10 pairs the two symbols were different. In the latter pairs, a string of symbols was paired with stars (*** or ****) to create non-identical pairs. The side on which the stars and the symbols appeared was counterbalanced across trials. For the picture baseline condition, scrambled pictures were created, using Adobe ImageReady (version 3.0). 26 pairs had two identical scrambled images and 10 pairs had non-identical scrambled images. In the latter pairs, the same picture was scrambled in two different resolutions, creating visually different images. Appendix 6 lists the words and pictures used in the rhyme judgement tasks and gives examples for the scrambled pictures and meaningless symbols used in the baseline conditions.

In addition, four practice blocks were used. In the practice session which was completed outside the scanner, participants saw different words/pictures than during the scan. The practice blocks were shorter, containing 20 pairs, 10 of which were

[13] The ratio between 'yes' (rhymes) and 'no' (does not rhyme) trials was set to 7:3, since trials in which the correct answer is 'yes' are more informative than those in which the correct answer is 'no'. In order to answer a 'yes' trial correctly, subjects have to produce the two words correctly. However, on trials where the correct answer is 'no', participants can produce many different errors, all resulting in the answer 'no'. As a result, 'yes' trials are more controlled and it was therefore desirable to have more 'yes' than 'no' trials in the experiment.

rhyming/matching pairs, and 10 were non-rhyming/non-matching pairs. For details of the different conditions see table 6.1.

Table 6.1. Conditions of the fMRI study.

Task	**Presentation mode**	**Rhyming word pair (YES)**	Non-rhyming word pair (NO)
1	Words	chair-bear	chair-fly
2	Pictures	pear-square	pear-door
3	Base line for words	Same meaningless symbols	Different meaningless symbols
4	Base line for pictures	Same scrambled pictures	Different scrambled picture

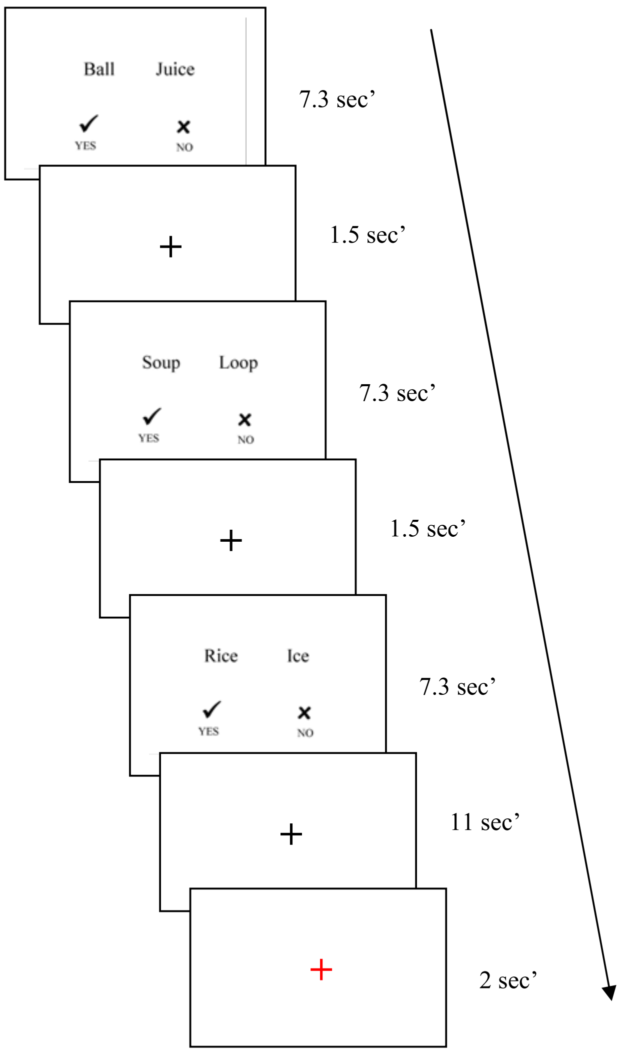
The time length of the inter-stimulus-interval (ISI) was determined based on pilot work in which patients with aphasia performed the tasks with no time limit, but were asked to do it as fast and as accurately as possible. It was found that the majority of trials were successfully answered in less than 7 seconds.

2.3 Procedure

First participants practiced the tasks outside the scanner. The task used in the training session was identical to the one used during the scan, except that different words and pictures were used. After completion of the training task and before entering the scanner, participants were shown the pictures and asked to name them. Naming errors were documented and corrected. When a naming error occurred the correct name was given, and participants were asked to name those pictures again until named correctly. This procedure was employed by Levelt and colleagues (for example, in Levelt, Praamstra et al. 1998), in their picture naming studies and it has a few advantages. Firstly, it ensures that participants produce the desired words during the scan. Secondly, naming the pictures in a relaxed environment, before entering the scanner, allows patients in particular to feel more confident with the task and could, therefore, improve their performance. Lastly, a previous study has suggested that participants who were shown the words prior to scanning with a picture-naming task had significantly greater brain signal than participants who were not exposed to the words beforehand (Jancke and Shah 2004).

In the scanner, participants performed two rhyme judgement tasks and two visual similarity tasks in four separate blocks. They were able to rest between blocks. In each trial of the rhyme judgement participants were presented with either two pictures or two words and had to indicate whether the words rhyme or not, by pressing one of two buttons, with their left hand (which is the non-paretic hand for the patients). In the baseline condition participants were asked to indicate whether two images were identical or not. In each trial, the words ‘yes’ and ‘no’, together with a ‘v’ and an ‘x’, respectively, appeared at the bottom of the screen to remind participants that the left button corresponds to ‘yes’ and the right button corresponds to ‘no’. Figure 6.1 shows the order of the stimuli presentation.

Figure 6.1. Order of stimuli as it appeared in the rhyme judgement experiment.



The stimuli was presented in blocks of similar stimulus, to maximise design efficiency (Price, Crinion et al. 2006), and because pilot work had shown that patients perform the tasks better when they were not required to alternate between tasks. The order of blocks was randomised and counterbalanced between participants. The order of yes/no answers was pseudo-randomised, making sure that the same word/picture did not repeat in two consecutive trials. The side on which each word appeared was counterbalanced between blocks. Images in the baseline tasks appeared equally on the left and right side of the screen. Stimuli were projected onto a back screen with a resolution of 1024x768 pixels.

2.4 Data Acquisition

Imaging was performed using a 3T Siemens (Erlangen, Germany) Allegra MRI scanner at the Wolfson Brain Imaging Centre, Cambridge. In each of the four imaging runs, 234 whole-brain functional T2*-weighted echoplanar images (EPI) (slice thickness: 3.75 mm, 32 axial slices, sequence: interleaved ascending, repetition time (TR): 2 sec, echo time (TE): 30 msec, flip angle: 78°; matrix: 64x64, field of view (FOV): 192x192 mm) were acquired. The first six volumes of each run were treated as dummy pulses and were discarded to allow for T1 equilibrium effects. In addition, a magnetization-prepared rapid-acquisition gradient echo (MPRAGE) scan was acquired (TR: 2.3 sec, TE: 2.98, FOV: 240x256 mm, sagittal plane; slice thickness: 1 mm; 176 slices).

2.5 Carotid Assessment

All patients who participated in the fMRI study had a carotid vascular assessment. The purpose of this assessment is to exclude patients with severe blood flow abnormality, which might lead to an abnormal haemodynamic response function (HRF). This assessment was performed by the medical staff as part of the routine treatment upon hospital admission, and included a Carotid Doppler with breath holding for neck, and in case of more than 50% stenosis, a CTA/MRA of the head and neck.

2.6 Data Analysis

2.6.1 Behavioural Data Analysis

Two main variables were measured in the study: response time (RT) and percentage of correct responses. Behavioural data was analysed using SPSS 16.0 (SPSS Inc., 1989-2007). Very fast responses (RT $\leq$ mean minus three standard deviations) were considered outliers and if found, excluded from further analysis. This is because such responses are unlikely to be a response to the trial but rather an accidental button press or a late response to a previous trial. If the overall performance was at chance level the participant's data was excluded.

2.6.2 fMRI Data Pre-processing and Analysis

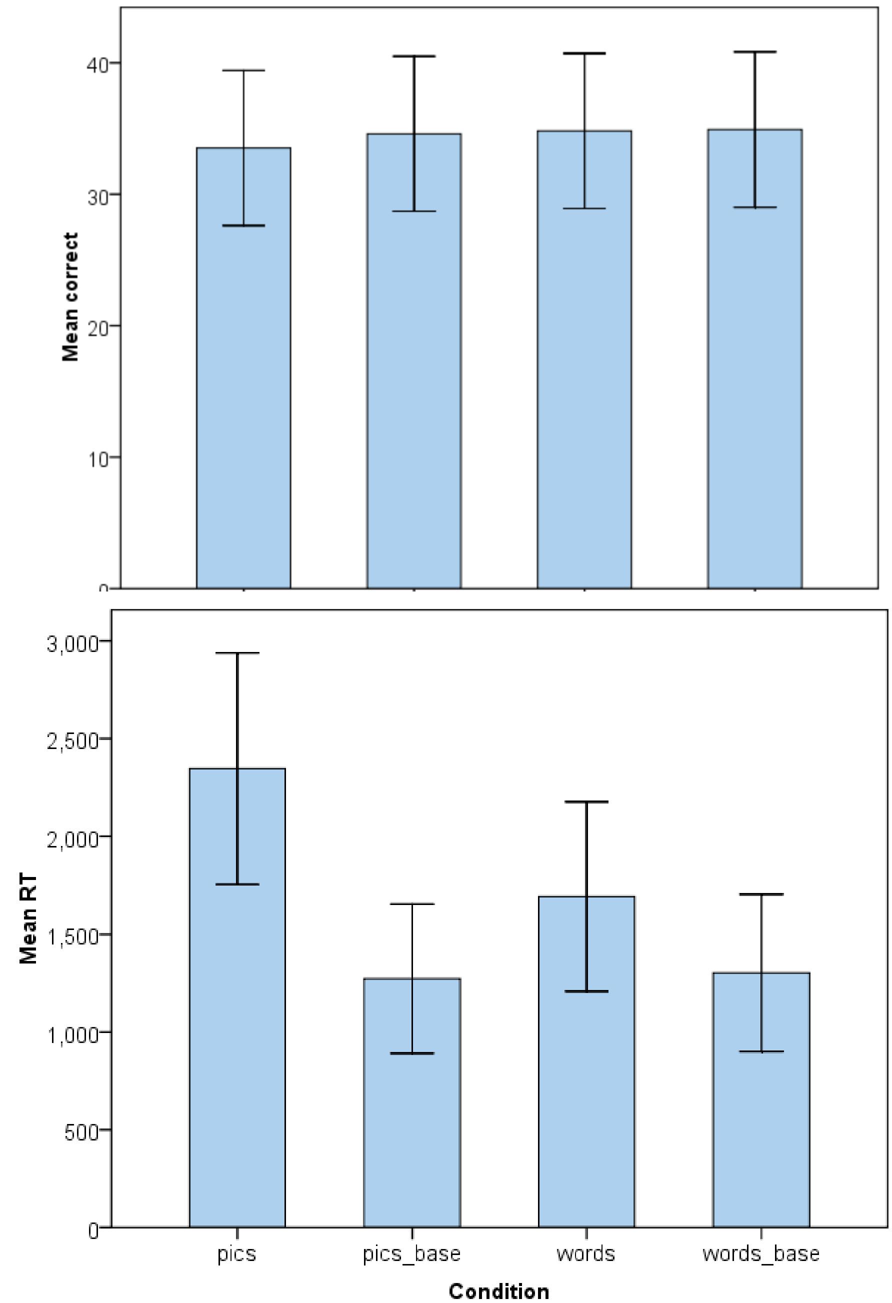
fMRI data was pre-processed using the Statistical Parametric Mapping software (SPM8, Wellcome Department of Cognitive Neurology, UCL) implemented in the Matlab (2006b) environment. Motion correction was performed using *Realign*, by first registering all images to the first image (after excluding the 6 dummy scans), and then registering all to the mean. Functional scans and structural MPRAGE scans were coregistered to each other using the *Coreg* function. Normalisation was conducted using the unified segmentation-normalisation procedure, in which the MPRAGE image is segmented into grey matter, white matter and CSF using *Segment*, and registered into the standard Montreal Neurological Institute (MNI) space template available in SPM8. The normalisation parameters were then applied to the functional images using the *Normalise* procedure. Lastly, data were spatially smoothed using an 8 mm full-width-at-half-maximum Gaussian kernel to reduce noise. BOLD response was convolved with the canonical HRF and significant group activation was determined at a threshold of $p<0.001$, uncorrected. Single subject activations are reported with a threshold of $p<0.01$ corrected for multiple comparisons using Family Wise Error (FWE). Correct and incorrect responses were modelled at the first level, as separate conditions. Motion parameters were added as multiple regressors. Scans of control subjects were analysed as a group using a whole brain analysis. Acknowledging the heterogeneity between patients (Benson and Ardila 1996), patients data was analysed individually. ROI analyses examined correlations between activation and behaviour or grey matter density. ROIs were defined using independent data as described below. Group activation maps are displayed on the single subject (Collin) template brain provided with SPM8. Patients' activation maps are displayed on both the single subject template brain and on the patient's own normalised T1 scan.

3. Results

3.1 Behavioural Results

No data was excluded based on fast RT or high level of error rate. The group of patients and older controls did not significantly differ in age (one-way ANOVA, $F_{23,1}$=0.29, p=0.59) or years of education (one-way ANOVA, $F_{23,1}$=0.004, p=0.95). Behavioural data significantly deviated from normality (Kolmogorov–Smirnov test, p>0.05 for all data). Therefore, non-parametric tests were subsequently used. The three groups (young controls, older controls and patients) did not differ in RT or error rate for any of the four tasks (Kruskal-Wallis test, p>0.1 for all tests). Therefore, any difference seen in brain activation cannot be attributed to overall differences in task performance (Christoff, Prabhakaran et al. 2001). Comparing the different conditions, it was found that both RT and the number of correct responses significantly differed between conditions (RT: Chi squared=72.2, p<0.001; number of correct responses: Chi squared=33.3, p<0.001). Planned comparisons showed that, as expected, RTs were longer for the picture condition compared with the word condition (Wilcoxon Signed Ranks Test, z=-5.3, p<0.001). The word condition had longer RTs than both baseline conditions (Wilcoxon Signed Ranks Test, z<-4.0, p<0.001 for both comparisons), and the baseline conditions did not differ from each other (Wilcoxon Signed Ranks Test, z=-0.35, p=0.72). Looking at error rates, it was found that the picture condition had more errors (less correct responses) than the word condition (Wilcoxon Signed Ranks Test, z=-3.7, p<0.001), and more errors than its baseline condition (Wilcoxon Signed Ranks Test, z=-2.6, p=0.01). No other comparisons were significant. Together, these behavioural data confirm that the picture condition was the most difficult, followed by the word condition, followed by the two baseline conditions which did not differ from each other. Behavioural data is presented in figure 6.2.

Figure 6.2. Mean number of correct answers (top) and mean RT (bottom), for each of the four conditions in the fMRI study (pics=picture rhyming task, words=words rhyming task, pics base=baseline condition for the picture task, word base=baseline condition for the word task). Error bars represent STD.



3.2 fMRI Results

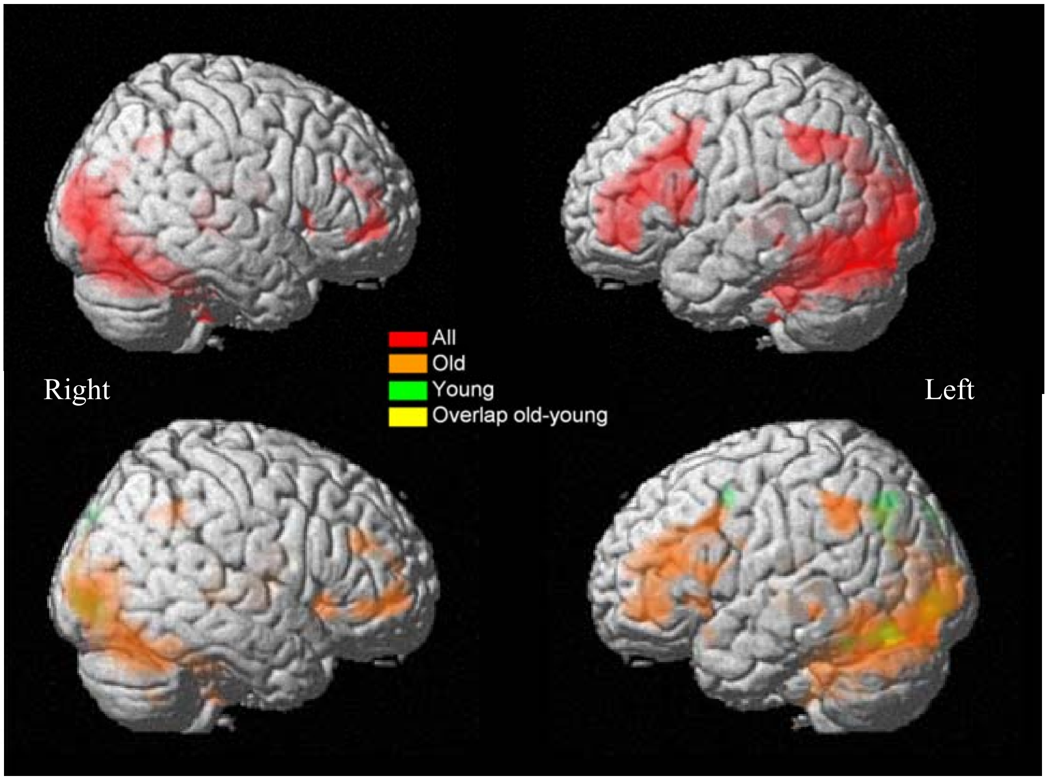
3.2.1 Inner Speech in Healthy Volunteers and Ageing

3.2.1.1 Word Rhyme Judgement

For the entire group of control participants (n=31) the word task (word>rest) produced activation in three large clusters: occipital-temporal, inferior parietal and frontal (figure 6.3, top, table 6.2). In the first area, activation was found in bilateral occipital regions, fusiform gyrus and superior parts of the cerebellum. The second activated region, in the left inferior parietal lobe, included the posterior inferior part of the supramarginal gyrus and the inferior angular gyrus. Lastly, frontal activation was seen in the left IFG (pars opercularis; BA44 and pars triangularis; BA45) and the inferior part of the MFG, including the DLPFC (BA46) and the premotor region (BA6). In the right hemisphere, frontal activation was restricted to prefrontal regions including the DLPFC (BA46) ($p<0.001$, uncorrected).

The group of older controls (n=19) showed widespread activation, very similar to that of the entire control group (figure 6.3, bottom panel, orange). The younger controls (n=12) showed activation only in the bilateral occipital regions, left fusiform gyrus and left angular/supramarginal gyrus. Frontal activation was absent, except for a small region of activation in the hand area of the motor strip (figure 6.3, bottom panel, green) ($p<0.001$, uncorrected).

Figure 6.3. Brain activation during the word rhyme judgement task (word rhyme>rest), for the entire control group (top, red) and for the older (bottom, orange), younger (bottom, green) and both (bottom, yellow) controls (p<0.001, uncorrected).



Comparing the older and younger controls directly, it was found that while the young controls did not show any activation which was significantly stronger than the older controls, the opposite effect (older>young) revealed significantly greater activation that was widely distributed but particularly noticeable in the right hemisphere. Greater activation was found in the midline and left cerebellum, right insula, right IFG (pars opercularis; BA44), and bilateral DLPFC (BA46 on the left, and BA9 on the right) (p<0.001, uncorrected).

Comparing the word task to its baseline, significant activation for all controls was found in the left hemisphere in the angular gyrus, ITG/fusiform gyrus and IFG (pars triangularis; BA45). The young controls showed activation only in the left angular gyrus, while the older controls showed activation in the left IFG (pars triangularis and orbitalis). Comparing the two control groups directly (older>young; young>older) did not yield any significant differences in activation (p<0.001, uncorrected).

Table 6.2. Regional activation for words>rest, all control participants. t values are listed only for the peak coordinate in every cluster. x, y, z coordinates are given in MNI space. Only clusters with a minimum of 20 voxels are reported.

Brain Region	Brodmann Area	x	y	z	Peak t value
Occipital lobes / Fusiform Gyrus					
R Calcarine	18	12	-96	0	9.51
L Fusiform Gyrus	19/37	-42	-68	-16	
L Fusiform Gyrus	37	-48	-58	-16	
Parietal Lobes					
R Angular Gyrus	40/7	26	-60	40	4.61
R Inferior Parietal	40	44	-48	44	3.76
L Inferior Parietal	39/40	-46	-46	38	5.78
Temporal lobes					
L Mid Middle Temporal Gyrus	21	-62	-34	-8	3.79
L Inferior Temporal Sulcus	20/21	-56	-38	-12	
Frontal Lobes					
R Mid Frontal	46/47	46	54	0	5.23
R Mid Frontal	10/46	40	58	-4	
R Inferior Frontal, pars orbitalis	47	42	48	-8	
R Inferior Frontal, pars opercularis	45	54	20	-2	4.11
L Inferior Frontal, pars opercularis	44	-54	16	34	5.96
L Inferior Frontal, pars triangularis	45	-42	30	18	
L Inferior Frontal, pars triangularis	45/46/47	-50	44	2	
Subcortical Regions					
R Hippocampus	27	28	-30	4	6.13
		24	-28	12	
		28	-20	-2	
R Putamen		24	-4	14	3.72
L Hippocampus	27	-26	-32	4	5.23
		-26	-20	-2	
L Thalamus		-20	-24	16	
Cerebellum and Brain Stem					
Pons		0	-30	-46	4.49
Pons		14	-32	-40	

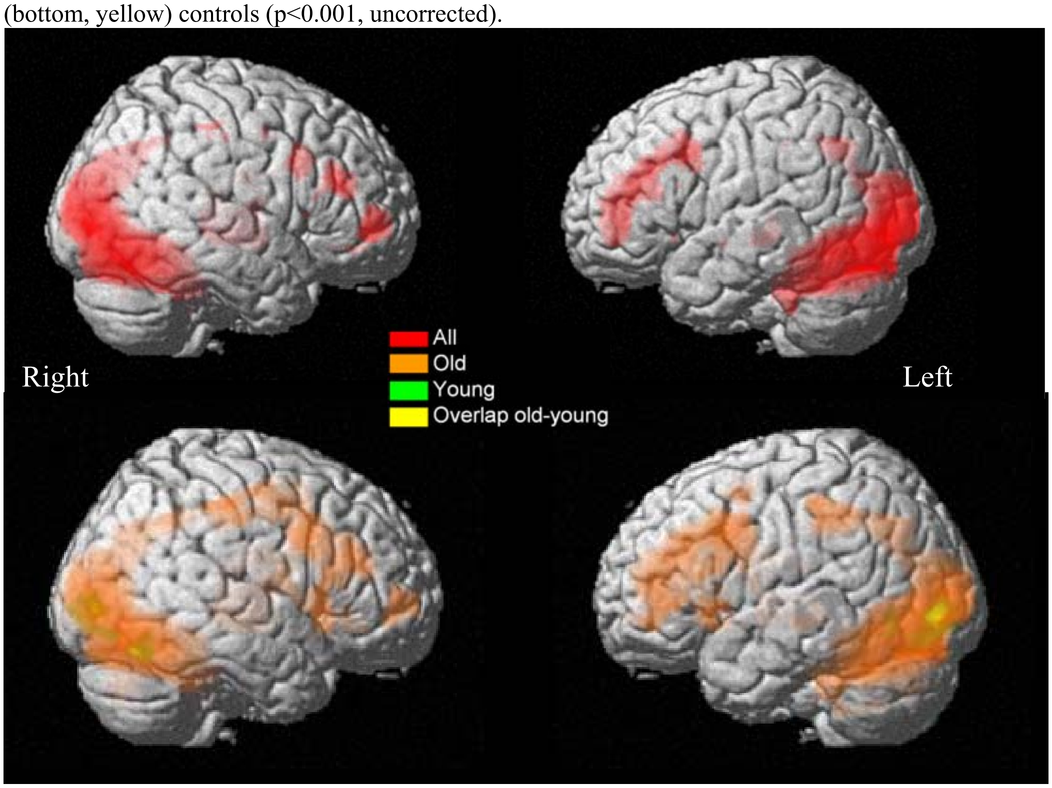
3.2.1.2 Picture Rhyme Judgement

In the picture rhyme judgement task healthy controls showed activation (pictures>rest) in three main clusters. The most posterior region of activation included bilateral occipital areas, extending anteriorly into the posterior part of the angular gyrus and into the fusiform gyrus. Secondly, activation was found in bilateral hippocampal regions. Lastly, frontal activation was found in an area extending along the MFG (from the pre-central gyrus to the DLPFC) and the IFG (pars triangularis; BA45). On the right, two smaller areas of frontal activation were found: in the most anterior part of the DLPFC (BA46) and in the superior part of the IFG, pars triangularis ($p<0.001$, uncorrected; figure 6.4, top, table 6.3).

The group of older controls (n=19) showed widespread activation, very similar to that of the entire control group (figure 6.4, bottom panel, orange). Additional activation was found in the bilateral insula and left IFG (pars opercularis; BA44). Activation in the group of the young controls appeared only in occipital and inferior temporal regions (figure 6.4, bottom panel, green) ($p<0.001$, uncorrected).

Comparing the two control groups directly, it was found that the older controls, compared to the young controls (older>young) showed significantly stronger activation in bilateral cerebellum, left pre-central gyrus, left supplementary motor area (BA6), right pre- and post-central gyrus and right IFG (pars triangularis and opercularis). No opposite effects (young>older) were found ($p<0.001$, uncorrected).

Figure 6.4. Brain activation during the picture rhyme judgement task (picture rhyme>rest), for the entire control group (top, red) and for the older (bottom, orange), young (bottom, green) and both (bottom, yellow) controls ($p<0.001$, uncorrected).



The picture rhyming task was compared to its baseline (picture rhyme>baseline). The control group showed small areas of activation in three left hemispheric regions: the inferior parietal cortex, IFG (pars triangularis) and ITG/fusiform gyrus. For the older controls, significant activation was found in the left IFG; pars triangularis (BA45) and pars orbitalis (BA47) and in the left fusiform gyrus. Young controls showed significant activation in the left angular-inferior parietal region. Comparing the two groups directly (young>older; older>young) did not reveal any significant activations (p<0.001, uncorrected).

Table 6.3. Regional activation for pictures>rest, all control participants. t values are listed only for the peak coordinate in every cluster. x, y, z coordinates are given in MNI space. Only clusters with a minimum of 20 voxels are reported.

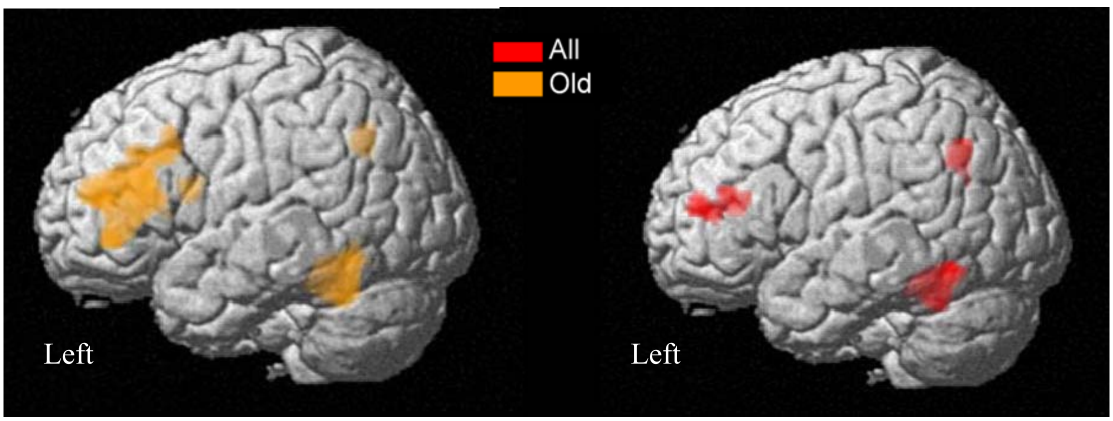
Brain Region	Brodmann Area	x	y	z	Peak t value
Occipital lobe / Fusiform Gyrus					
R Inferior Occipital Gyrus	19	36	-84	-2	
L Mid Occipital	19	-34	-86	2	10.54
L Lingual Gyrus	18	-20	-80	-14	
Parietal Lobe					
R Post-central Gyrus	3	52	-24	50	3.96
R Supramarginal Gyrus	2/40	50	-32	48	
L Inferior Parietal	7/40	-26	-58	40	4.54
Frontal Lobe					
R Mid Frontal	46	46	54	-2	6.94
		44	54	6	
R Inferior Frontal, pars triangularis	45	48	34	26	5.6
		58	24	16	
R Inferior Frontal, pars opercularis	44	54	16	34	4.43
R Middle Frontal Gyrus	6	26	-2	44	3.88
R orbitofrontal area	11/47	30	42	-12	3.82
L Frontal Mid	45/46	-48	40	20	5.01
L Inferior Frontal, pars triangularis	45	-46	46	12	
L Inferior Frontal, pars triangularis	44/45	-50	30	28	
Subcortical Regions					
R Hippocampus	27/20/37	24	-26	0	7.99
R Putamen		24	-10	8	
R Pallidum		28	-6	-8	
R Caudate		20	-10	24	3.84
R Caudate		20	-2	26	
L Hippocampus	27/20/37	-26	-28	-4	6.66

3.2.1.3 Cognitive Conjunction Analysis

A conjunction analysis (Price and Friston 1997) explored the neural correlates which support inner speech, over and above presentation mode: That is, areas which are related to inner speech in both the word reading and the picture naming tasks. Conjunction analysis identifies a main effect in the absence of interaction. In imaging terms, this means that the analysis reveals areas which are active in both tasks, to the same degree, and do not interact with the baseline tasks, i.e. they are not context specific. For example, an area which is active during inner speech but only if the task involves written material, would be considered context-specific, or showing an interaction effect, and will therefore not be significant in a conjunction analysis.

Control participants showed significant conjunction for the word and picture conditions (word>baseline AND pictures>baseline) in three left hemispheric regions: the angular gyrus, the fusiform gyrus and the IFG (pars triangularis; BA45) ($p<0.001$, uncorrected; figure 6.5, right). Older participants showed a similar pattern to the one described for the entire control group. In this group, the clusters were bigger, and in the case of the IFG included the pars opercularis (BA44) and pars orbitalis (BA47) ($p<0.001$, uncorrected, figure 6.5, left). Young controls did not show any significant activation for the conjunction of the word and picture tasks ($p<0.001$, uncorrected). These analyses did not yield any significant right hemispheric activation.

Figure 6.5. A conjunction analysis for the word and picture rhyming tasks ($p<0.001$, uncorrected), for the entire controls group (right, red) and for the older controls only (left, orange).



3.2.1.4 Premotor Activation

The following analysis examined the relation between premotor activation and performance (percentage of correct responses and RT). The VBM analysis (chapter 5) showed that rhyme judgement correlated with tissue density in the premotor cortex (BA6), over and above overt speech production. This correlation was found when analysing the data from the patients alone, as well as when analysing the data from all participants. The local maxima within the premotor area was found at coordinate (-32 , 8 , 46). A spherical region of interest (10 mm radius) was created around this coordinate. Activation in this region of interest (percent signal change associated with correct trials) in the word task was correlated with RT within the group of older participants ($r^2=0.538$, $p=0.009$) but not within the group of young participants ($r^2=0.054$, $p=0.434$). Activation was not correlated with error rate for any of the groups ($p<0.05$). Correlation between activation in this region of interest and performance on the picture task was not significant for either groups ($p>0.05$ for all).

3.2.1.5 Correlations with Performance and Grey Matter Density

As shown above, younger and older controls did not differ in their performance on any of the tasks, in either percentage of correct responses or response time (Kruskal-Wallis test, $p>0.1$ for all tests). Compared to young controls, older controls showed stronger activation in the right IFG in both the word and the picture rhyming task. To explore the functional role of this right hemispheric activation, correlations between activation and behaviour or grey matter density were examined.

Comparing the older and young controls (older>young) in the word task, right IFG activation was found at a local maxima (40, 36, -2). This local maxima is in the pars orbitalis; BA47. A spherical region of interest (10 mm radius) was created to encompass this local maxima (central coordinate: 33, 37, -5). Within the group of older controls, activation in this region of interest (percent signal change associated with correct trials) showed a small, but yet significant correlation with RT ($r^2=0.42$, $p=0.039$). This was not significant for the young controls ($r^2=-0.052$, $p=0.436$). Correlation between activation and error rate was not significant for either groups ($p>0.05$ for both).

In the picture task, right IFG activation was significantly greater for the older controls (older>young) in the pars triangularis; BA45 (local maxima 36, 24, 16). A spherical region of interest (10 mm radius) was created around this local maxima. Correlation between activation in this region of interest and performance on the picture task was not significant for any of the groups ($p>0.05$ for all).

To examine the relation between structure and function, negative correlations between activation (percent signal change associated with correct trials) and grey matter density were calculated, within the cohort of older controls. It was found that activation in the right IFG during the word rhyming task did not significantly correlate with grey matter density in any brain regions ($p<0.0001$, uncorrected). Right IFG activation during the picture rhyme task was negatively correlated with grey matter density in the right cerebellum ($p<0.0001$, uncorrected).

3.2.1.6 Interaction between Presentation Mode and Ageing Effects

It was hypothesised that the effects of ageing will be more substantial in the picture condition than in the word condition. To test this hypothesis the interaction between age (older vs. young) and condition (pictures vs. words) was examined. The analysis examined areas which show ageing effects (older>young), and within these areas, looked for regions which show a condition effect (words>pictures or pictures>words). No areas showed an effect of words>pictures (p<0.001, uncorrected). This means that within the areas for which older controls showed overall more activation than younger controls (significant across conditions), no areas showed greater activation for words over pictures. The opposite analysis yielded some significant regions (p<0.001, uncorrected). That is, within the areas which showed an ageing effect (older>young), some regions also showed a condition effect (pictures>words). This included the right mid fusiform and superior part of the right cerebellum, as well as similar but smaller cluster on the left (data not shown).

3.2.2. Inner Speech in Aphasia

Patients' activation maps are displayed as follows: on the top panel of each figure rendered activation maps are superimposed on the single subject normalised brain template provided in SPM8. Activation during the word rhyming task (word>rest) is represented in green and activation during the picture rhyming task (picture>rest) is represented in blue. This figure allows easy comparison to the results from the control group. In addition, activation maps are displayed on representative slices of the patient's individual normalised T1 scan. Activation from the word rhyming task is displayed on the right set of T1 slices and activation from the picture rhyming task is on the left set of images. These figures allow a comparison of the activation maps to the patient's own lesion and anatomy. Lastly, activation is presented on the patient's normalised T1 in three sections: coronal, axial and sagittal, to facilitate anatomical localisation. Results are compared to those obtained from the older control group, who were matched to the patients on age, number of years of formal education and behavioural performance.

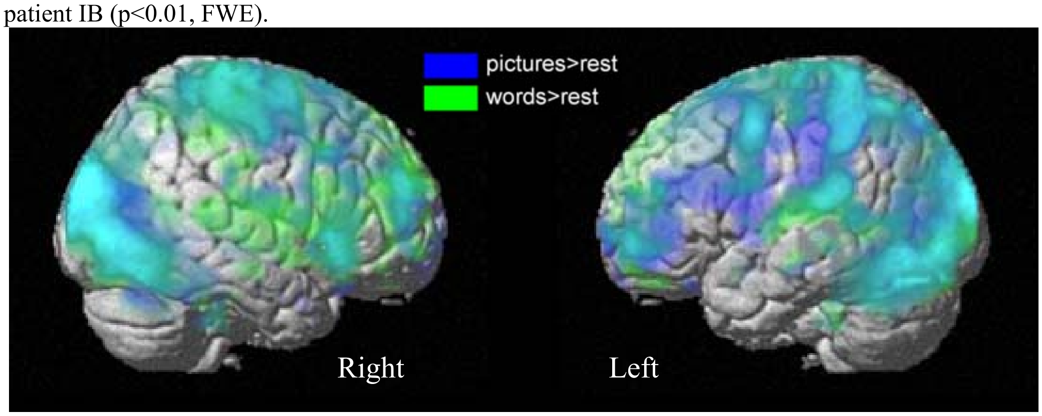
Patient **DB** had no clear significant activation. It should be noted that DB had excessive motion (more than 3 mm in all directions in the word task and more than 4 mm in all directions in the picture task) throughout the scans. Results from the other five patients are presented below.

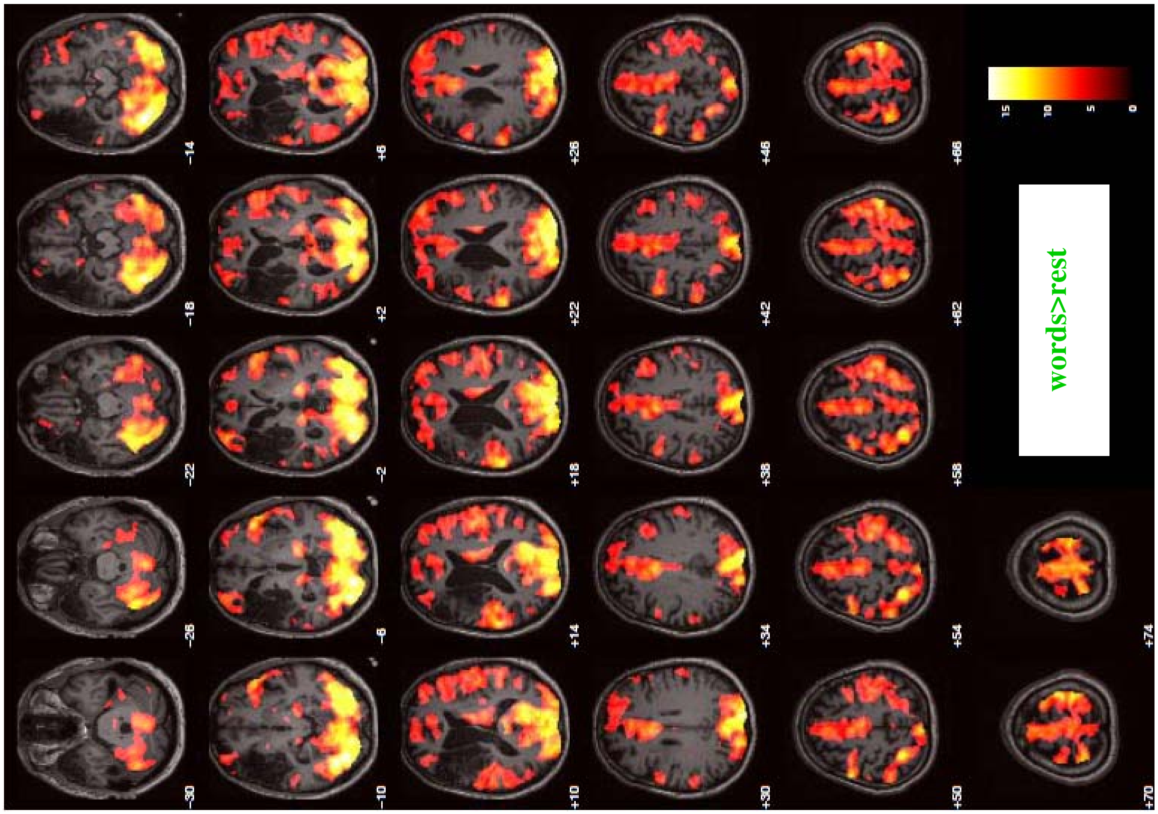
Patient **IB** had a large left perisylvian lesion involving the temporal pole, anterior parts of the middle and superior temporal gyri, insula, IFG and putamen. In the word task, patient IB showed widespread activation throughout the brain. This included clusters in bilateral occipital regions, extending into the fusiform and inferior temporal gyri, bilateral superior temporal, more widespread on the right, peri- and contra-lesional frontal activation, extending across the middle and inferior frontal gyri in the left, across the temporal lobe on the right, and into bilateral superior and inferior parietal regions ($p<0.01$, FWE; figures 6.6-6.7).

When comparing the word rhyme judgement to baseline, IB showed minor scattered areas of activation in occipital regions, as well as significant activation in the left temporo-parietal junction (STG and supramarginal), pre- and post-central gyrus bilaterally and SMA bilaterally ($p<0.01$, FWE). These results are similar to the activation obtained for the word task only (word>rest) though here, as expected, occipital activation was substantially reduced.

In the picture task, IB showed a pattern of activation which was similar to the one obtained during the word rhyme judgement task. Again, activation was extensive and covered most of the brain, with peak activations in bilateral occipital regions, left parietal cortex and left frontal lobe, around the lesion ($p<0.01$, FWE, figures 6.6-6.7). When comparing the picture rhyme judgement to its baseline condition, IB showed activation in most of the brain ($p<0.01$, FWE).

Figure 6.6. Brain activation during the word rhyming task (on this page, in green; on the next page, on the right), and during the picture rhyming task (on this page, in blue; on the next page, on the left) for patient IB ($p<0.01$, FWE).




words>rest
-30
-26
-22
-18
-14
-10
-6
-2
+2
+6
+10
+14
+18
+22
+26
+30
+34
+38
+42
+46
+50
+54
+58
+62
+66
+70
+74
15
10
5
0

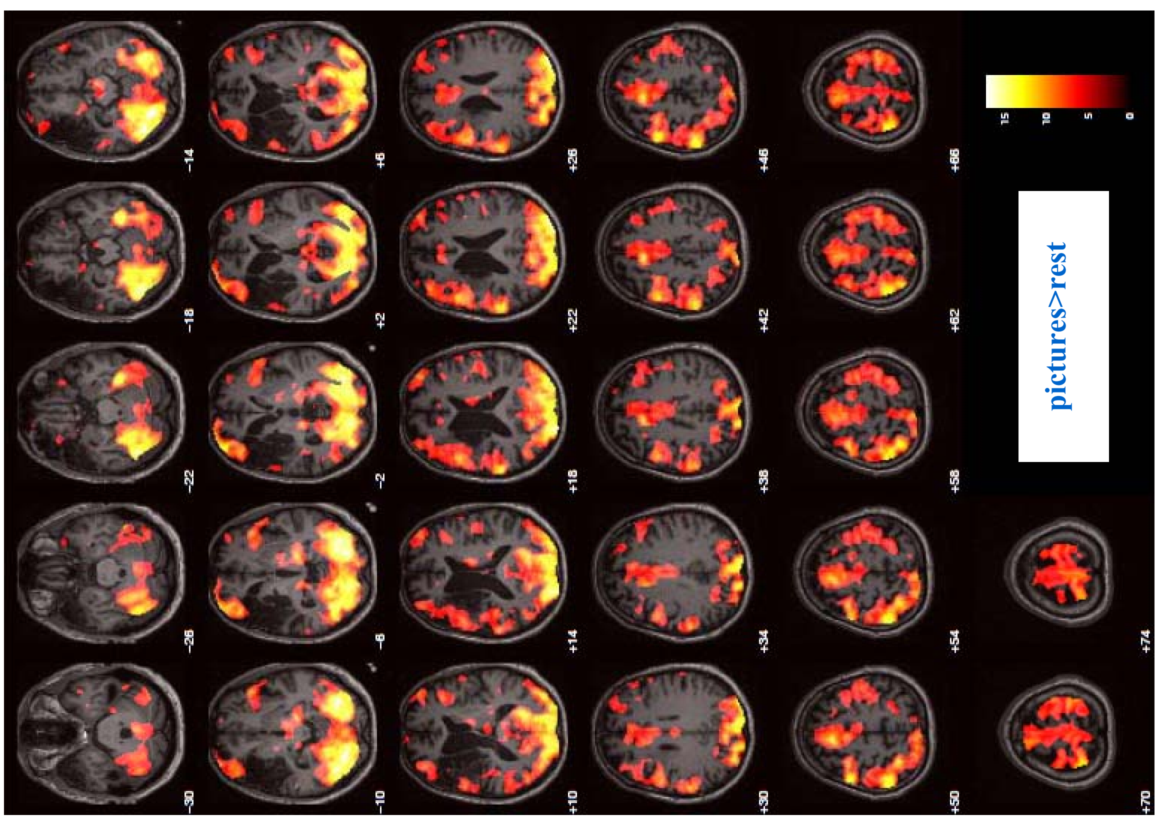
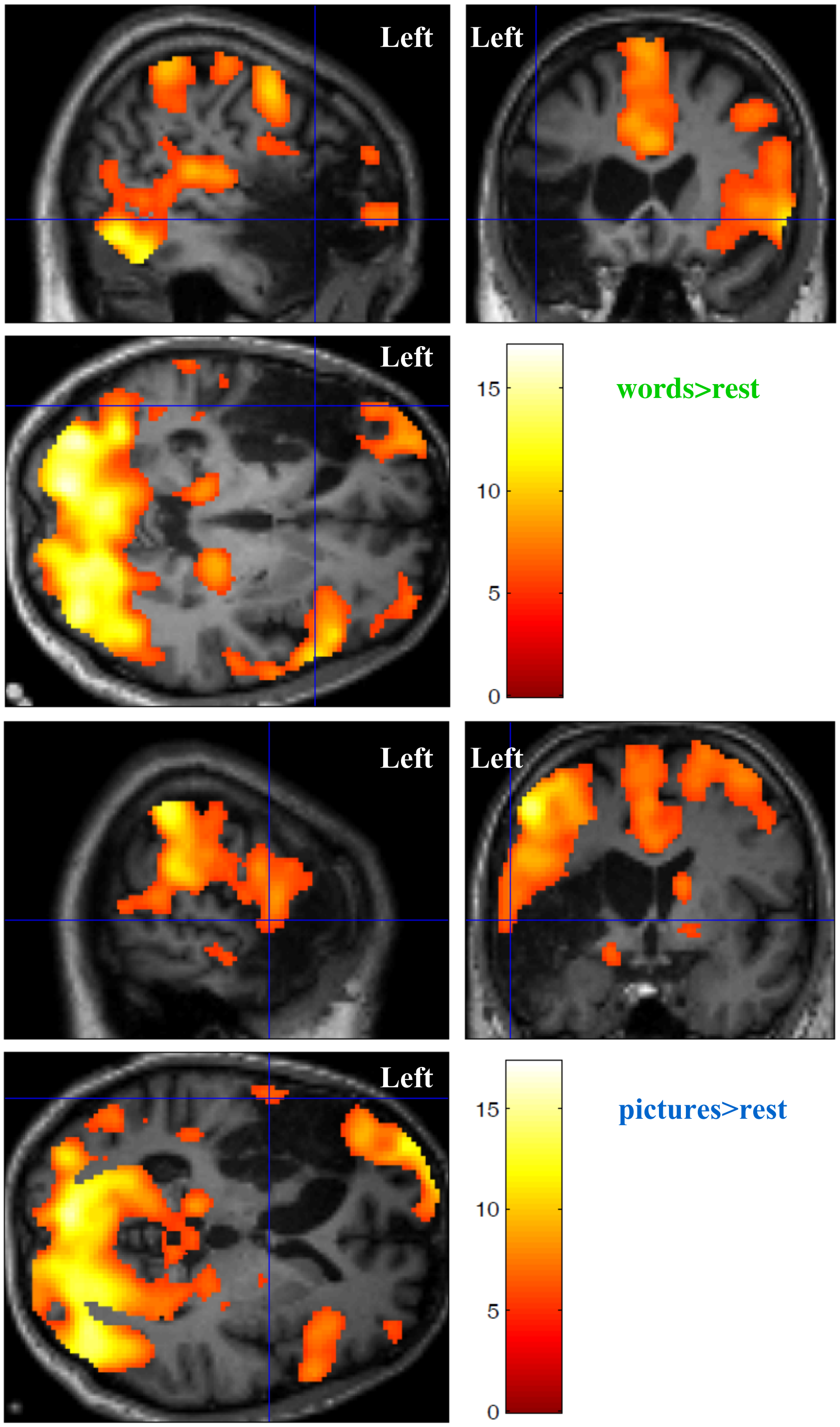

pictures>rest
-30
-26
-22
-18
-14
-10
-6
-2
+2
+6
+10
+14
+18
+22
+26
+30
+34
+38
+42
+46
+50
+54
+58
+62
+66
+70
+74
15
10
5
0

Figure 6.7. Sagittal, coronal and axial slices showing activation for words (words>rest; top) and pictures (pictures>rest; bottom) for patient IB.

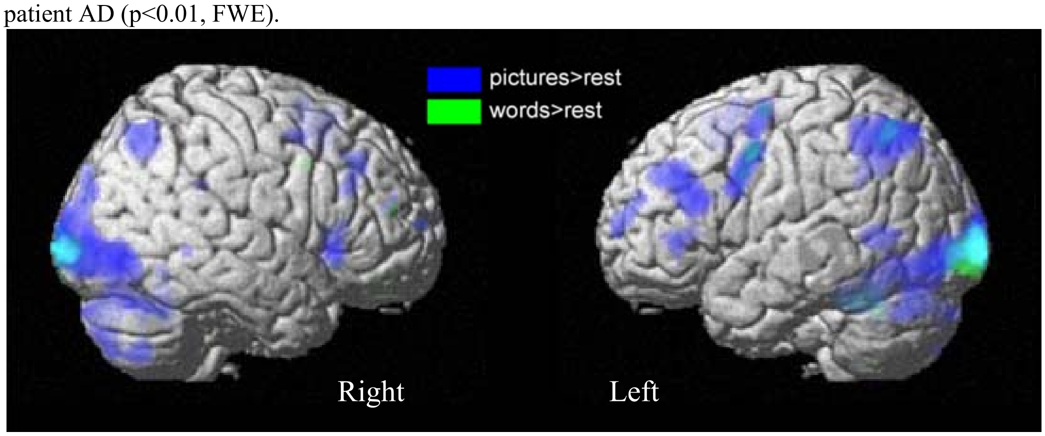


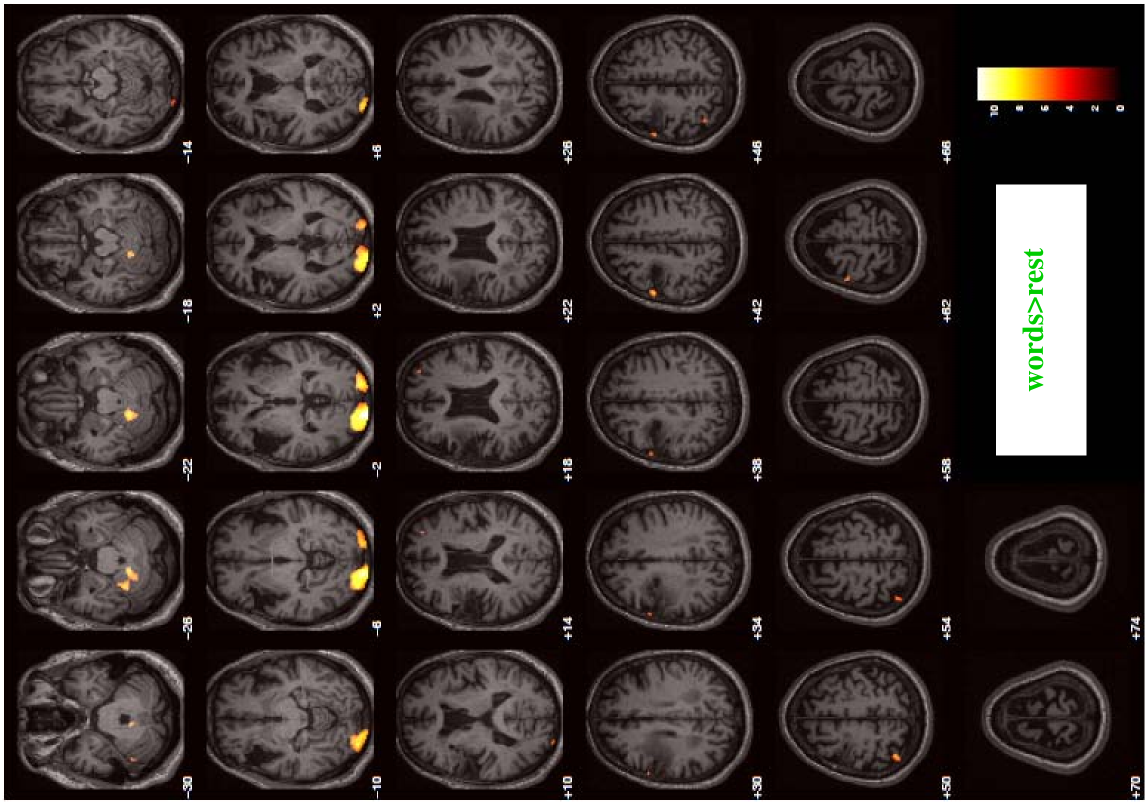
Patient **AD** had a left insular infarct which extends into the post-central gyrus. In the word task, patient AD showed activation in bilateral occipital regions, left cerebellum, superior part of the left angular gyrus, and in small peri-lesional regions in the left pre-central gyrus ($p<0.01$, FWE; figures 6.8-6.9). When comparing the word rhyme judgement to baseline, AD showed no significant activation ($p<0.01$, FWE).

In the picture task, AD showed activation in bilateral occipital areas, bilateral cerebellum (superior and inferior parts) and left fusiform gyrus. Left hemispheric activation included the angular gyrus, posterior MTG, pre-central gyrus, extending into the IFG; pars opercularis, along the anterior border of the lesion, SMA, middle and anterior (BA46) sections of the MFG, IFG (pars opercularis; BA44) and anterior insula. On the right, activation was found in the IFG (pars triangularis; BA45) and right anterior insula ($p<0.01$, FWE, figures 6.8-6.9).

When comparing the picture rhyme judgement to its baseline condition, AD showed activation in bilateral occipital areas (though more restricted to inferior occipital), and the left fusiform gyrus. Left hemispheric activation included the angular gyrus, posterior MTG, pre-central gyrus, extending into the IFG (pars opercularis), along the anterior border of the lesion, SMA, middle and anterior sections of the MFG, IFG (pars opercularis; BA44) and anterior insula. On the right, activation was found in the anterior insula.

Figure 6.8. Brain activation during the word rhyming task (on this page, in green; on the next page, on the right), and during the picture rhyming task (on this page, in blue; on the next page, on the left) for patient AD ($p<0.01$, FWE).




words>rest
−30
−26
−22
−18
−14
−10
−6
−2
+2
+6
+10
+14
+18
+22
+26
+30
+34
+38
+42
+46
+50
+54
+58
+62
+66
+70
+74
10
8
6
4
2
0

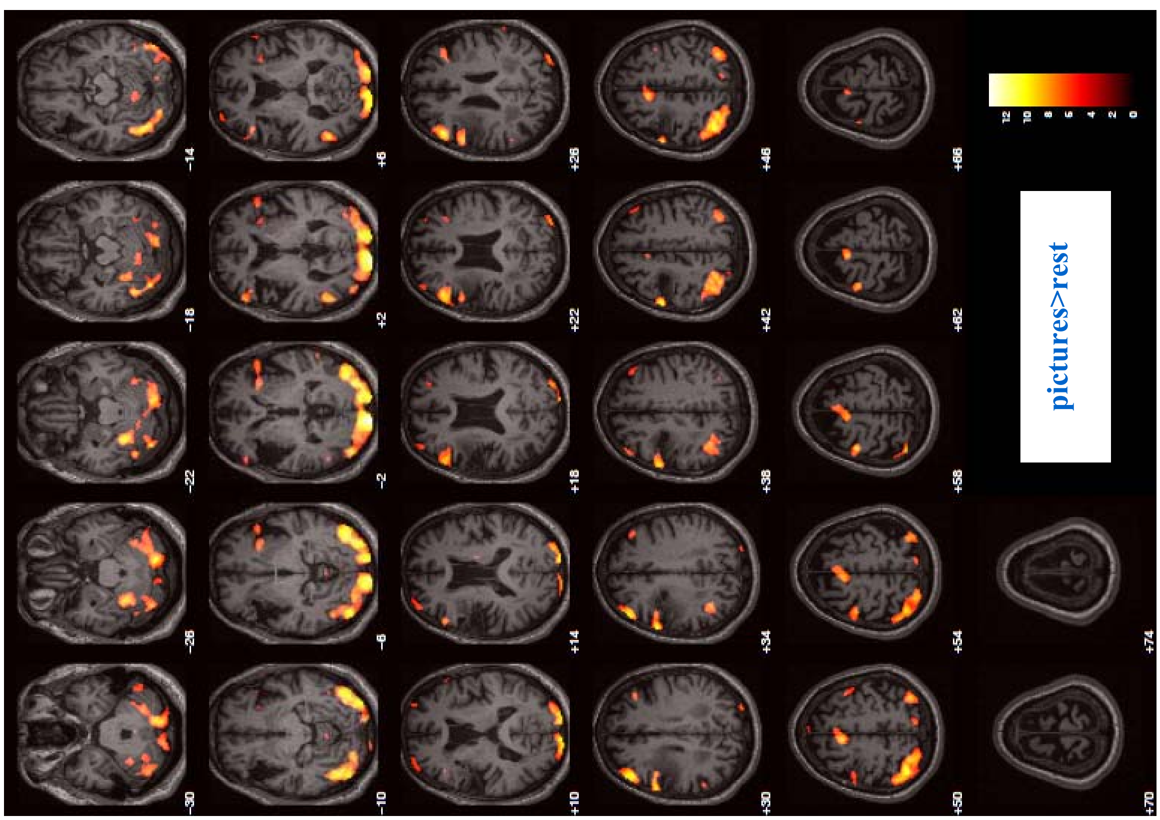
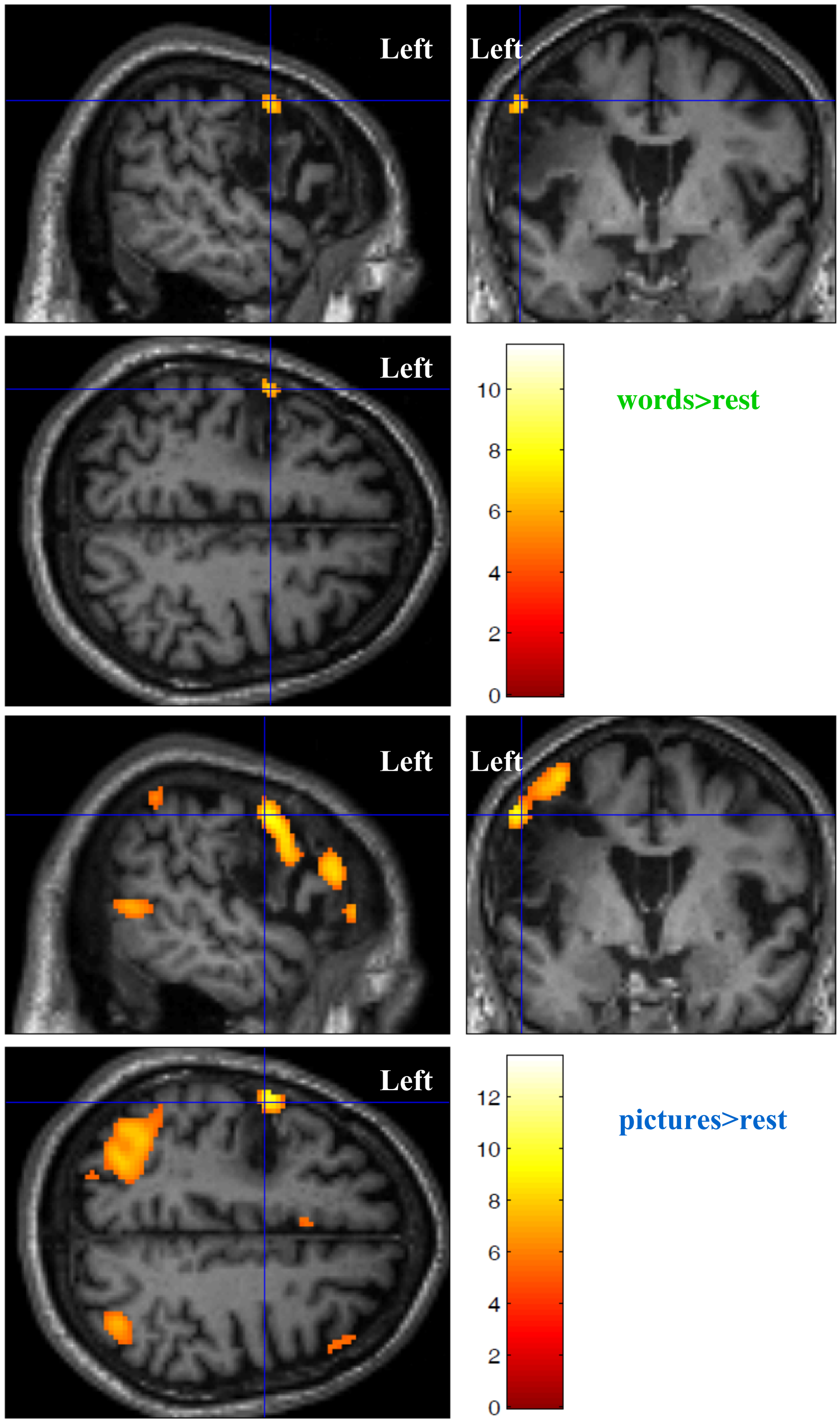

pictures>rest
−30
−26
−22
−18
−14
−10
−6
−2
+2
+6
+10
+14
+18
+22
+26
+30
+34
+38
+42
+46
+50
+54
+58
+62
+66
+70
+74
12
10
8
6
4
2
0

Figure 6.9. Sagittal, coronal and axial slices showing activation for words (words>rest; top) and pictures (pictures>rest; bottom) for patient AD.

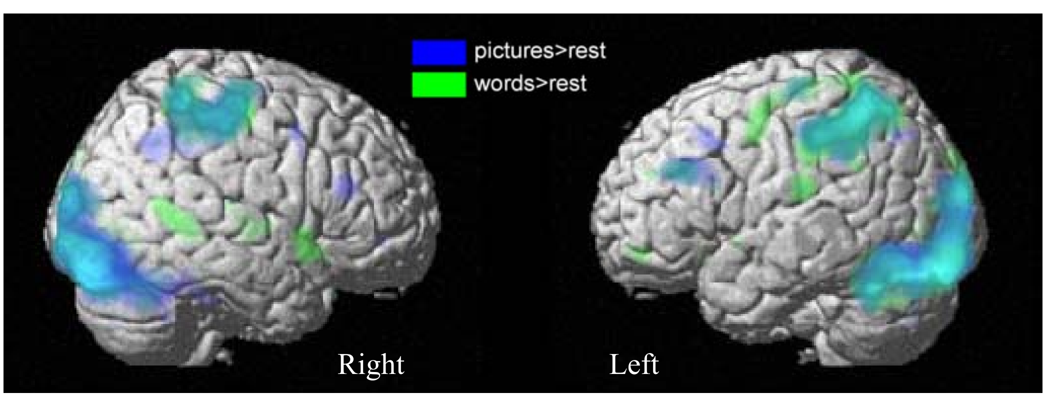


Patient **BA** had a lesion in the left inferior parietal lobe. In the word task, patient BA showed widespread activation in bilateral occipital and superior cerebellar regions, the left fusiform gyrus and the inferior temporal lobe. Parietal activation included left peri-lesional areas: the supramarginal and the superior parietal gyrus, as well as pre-central activation. In the right hemisphere, parietal activation included the anterior supramarginal gyrus and the inferior and superior parietal cortex, as well as the pre- and post-central gyri. In general, this right hemispheric parietal activation was more superior and anterior than the left hemispheric parietal activation. Right hemispheric temporal activation was found in the posterior MTG and the temporal pole. Frontal activation was found in the left anterior MFG ($p<0.01$, FWE; figures 6.10-6.11). When comparing the word rhyme judgement to baseline, BA showed no significant activation ($p<0.01$, FWE).

In the picture task, BA had an activation pattern which was very similar to that obtained above. Specifically, significant activation was found in bilateral occipital regions and bilateral superior cerebellum. On the left, a parietal region of activation included the inferior parietal lobe, supramarginal gyrus and angular gyrus. A similar region was found on the right hemisphere, with the difference that the right hemispheric activation extended anteriorly to the central sulcus. Frontal activation was found along the inferior frontal sulcus, and this was more extensive on the left than on the right ($p<0.01$, FWE, figure 6.10-6.11).

When comparing the picture rhyme judgement to its baseline condition, BA had no significant activation at the corrected threshold of $p<0.01$. However, with a lower threshold ($p<0.001$, uncorrected), significant activation was found in the right cerebellum, left angular gyrus, and left IFG (pars triangularis; BA45, and pars orbitalis; BA47).

Figure 6.10. Brain activation during the word rhyming task (on this page, in green; on the next page, on the right), and during the picture rhyming task (on this page, in blue; on the next page, on the left) for patient BA ($p<0.01$, FWE).



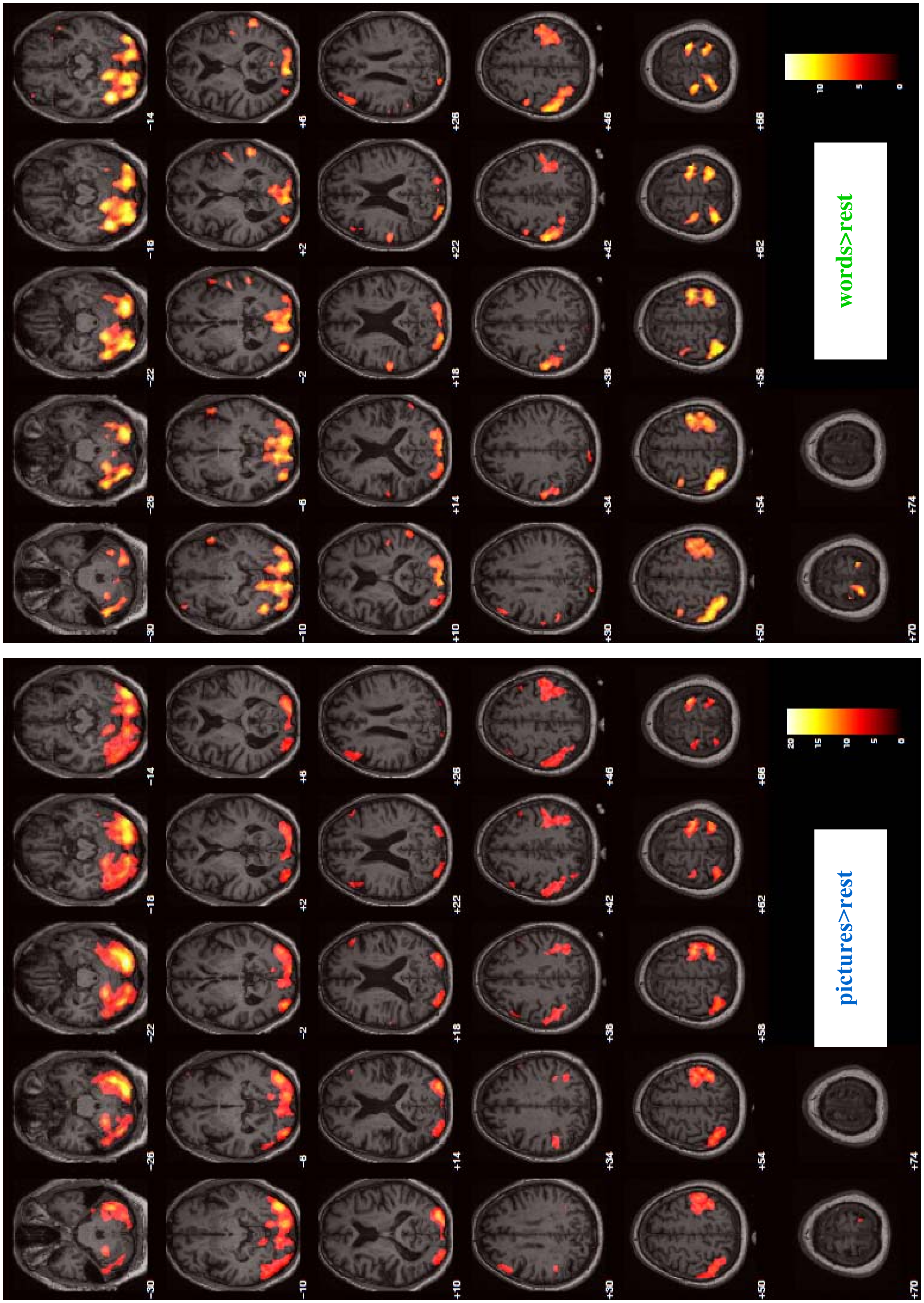
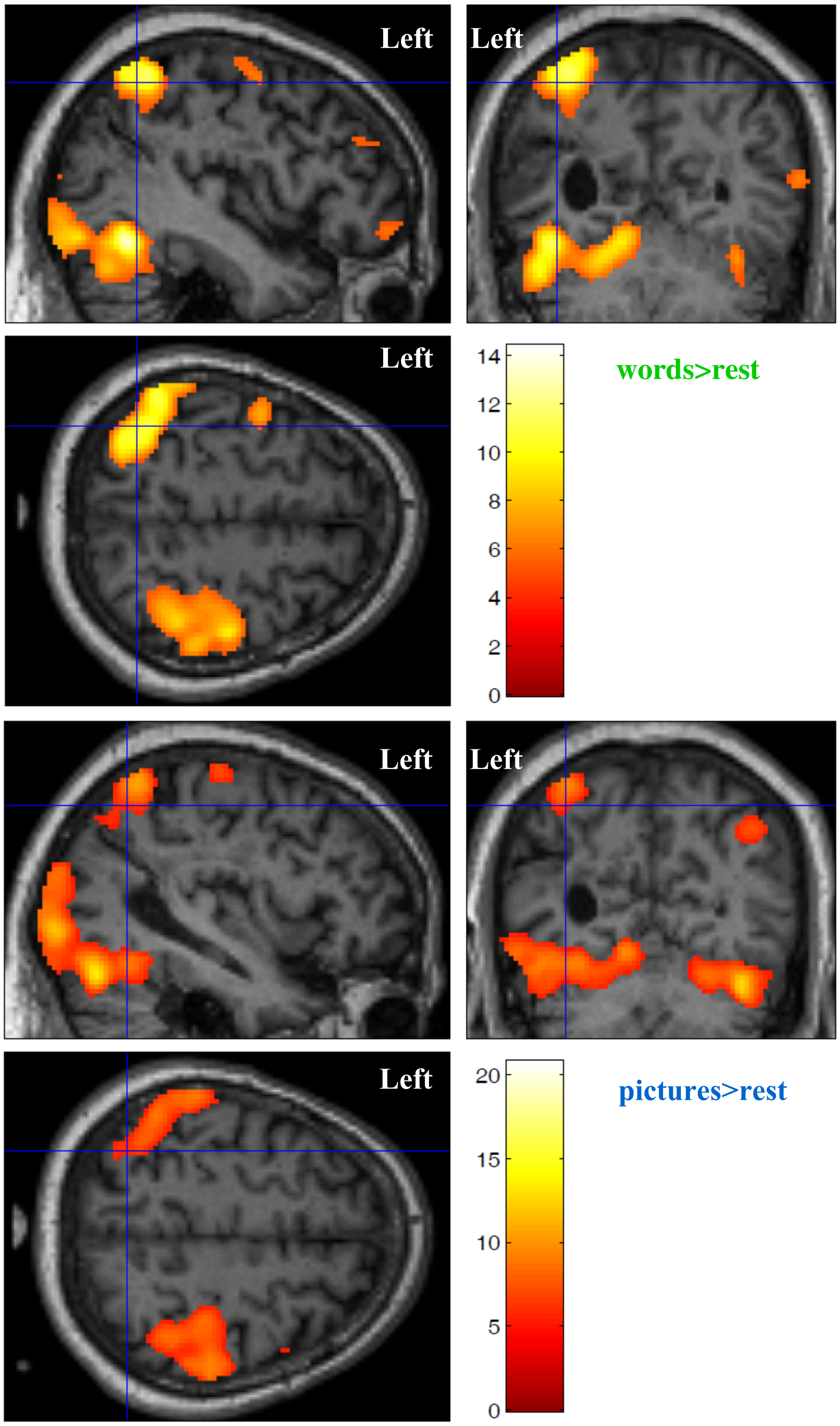

words>rest
pictures>rest

Figure 6.11. Sagittal, coronal and axial slices showing activation for words (words>rest; top) and pictures (pictures>rest; bottom) for patient BA.

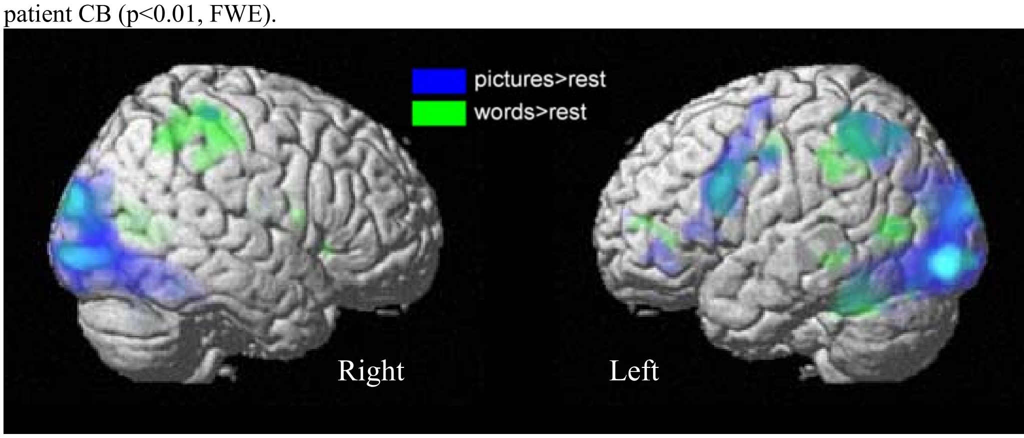


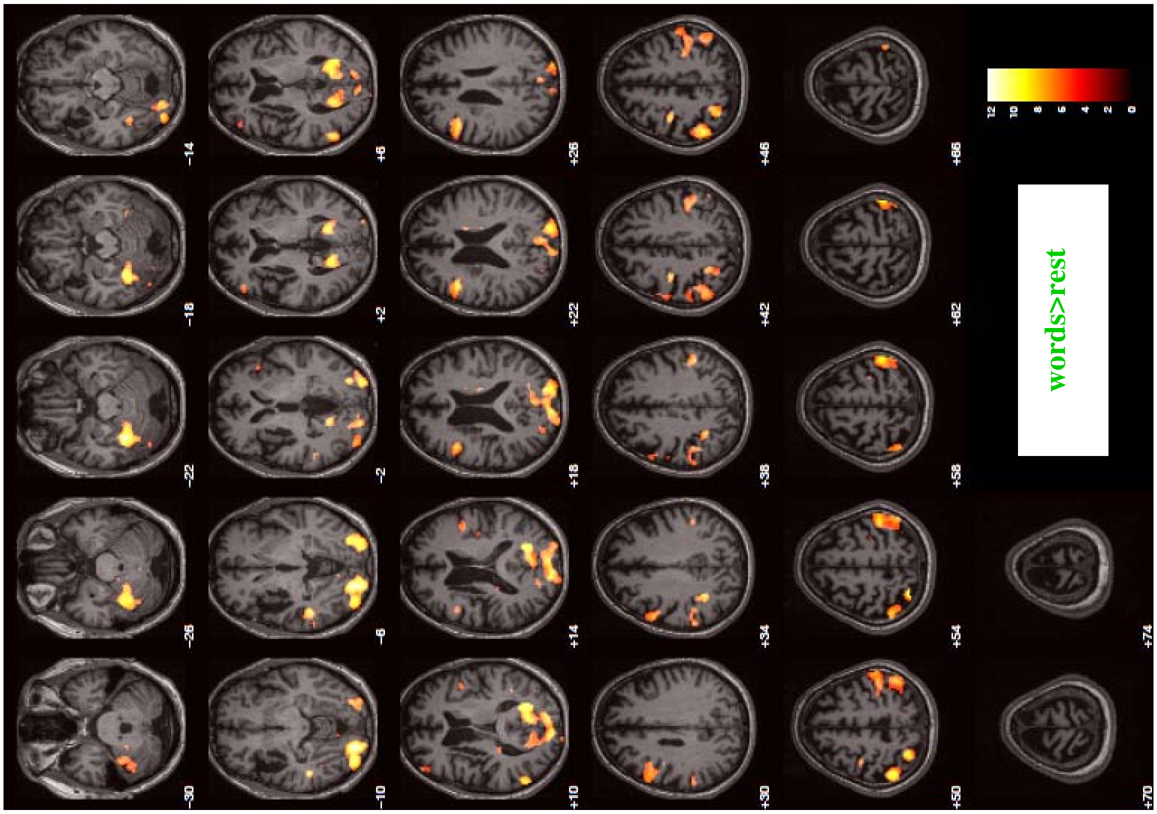
Patient **CB** had a left subcortical lesion involving the putamen, insula and subinsular white matter, as well as a lesion to the anterior temporal pole. In the word task, patient CB showed activation in bilateral occipital areas and the left fusiform gyrus. On the right hemisphere, activation was found in the inferior parietal region, extending into the superior part of the supramarginal gyrus, and in the IFG, pars opercularis (BA44). In the left hemisphere, activation was found in the inferior parietal cortex and supramarginal gyrus, posterior MTG, pre-central gyrus, IFG; pars opercularis (BA44) and DLPFC (BA46) ($p<0.01$, FWE; figures 6.12-6.13). When comparing the word rhyme judgement to baseline, CB showed no significant activation ($p<0.01$, FWE).

In the picture task, CB showed activation in bilateral occipital areas and the bilateral cerebellum. On the left, activation extended into the fusiform gyrus. Significant activation was also found in the left angular gyrus and inferior parietal lobe, and in two regions in the frontal lobe: one area extended across the premotor cortex (BA6), middle section of the MFG, and the IFG; pars opercularis (BA44), and the second covered the orbital prefrontal region (BA46) and IFG; pars orbitalis (BA47) ($p<0.01$, FWE, figures 6.12-6.13).

When comparing the picture rhyme judgement to its baseline condition, CB showed no significant activation at the corrected threshold of $p<0.01$, but with a lower threshold significant activation was found in bilateral occipital areas. On the left, activation was found in the fusiform gyrus, angular and inferior parietal lobe, and in two regions in the frontal lobe: one area extended across the MFG and IFG; pars opercularis (BA44) and triangularis (BA45), and into the insula, and the second covered the orbital prefrontal region (BA46) and IFG; pars orbitalis (BA47). Activation was also found in the right anterior insula ($p<0.001$, uncorrected).

Figure 6.12. Brain activation during the word rhyming task (on this page, in green; on the next page, on the right), and during the picture rhyming task (on this page, in blue; on the next page, on the left) for patient CB ($p<0.01$, FWE).




words>rest
-30
-26
-22
-18
-14
-10
-6
-2
+2
+6
+10
+14
+18
+22
+26
+30
+34
+38
+42
+46
+50
+54
+58
+62
+66
+70
+74
12
10
8
6
4
2
0

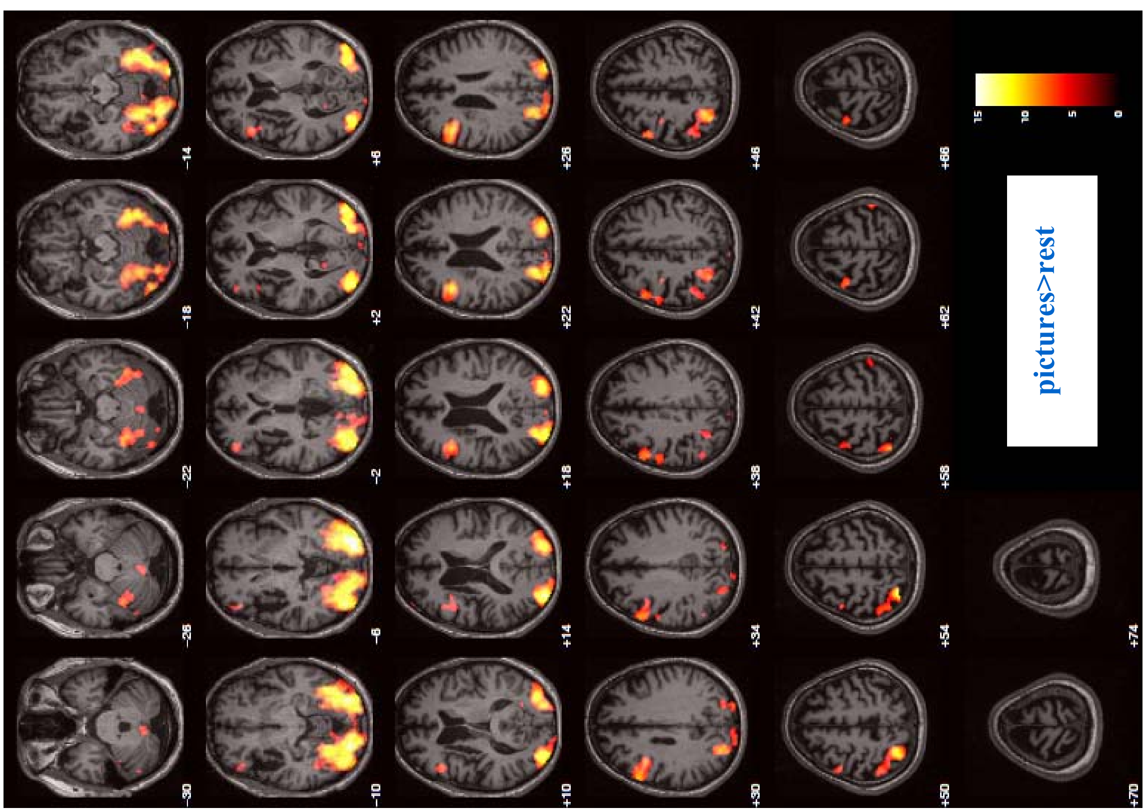
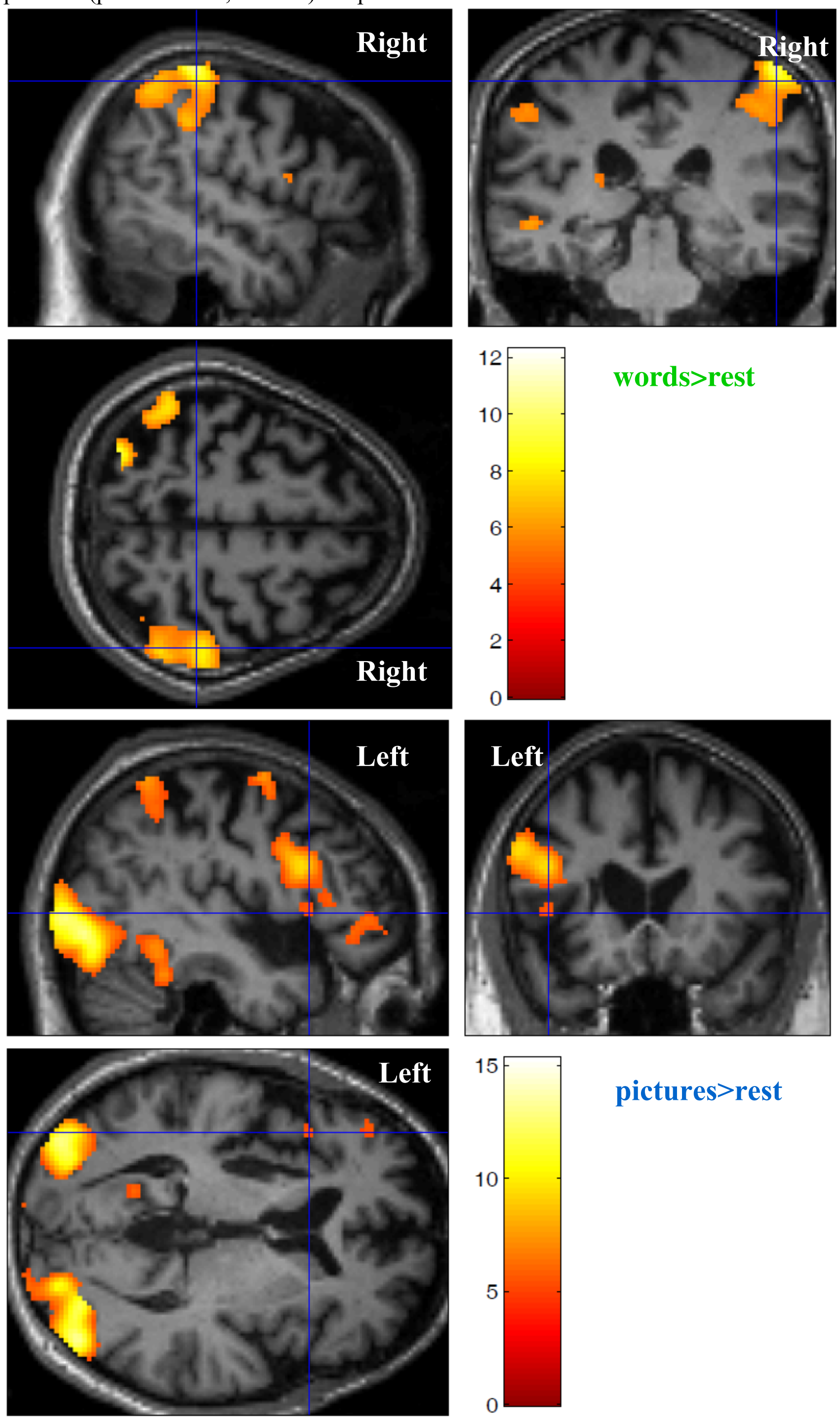

pictures>rest
-30
-26
-22
-18
-14
-10
-6
-2
+2
+6
+10
+14
+18
+22
+26
+30
+34
+38
+42
+46
+50
+54
+58
+62
+66
+70
+74
15
10
5
0

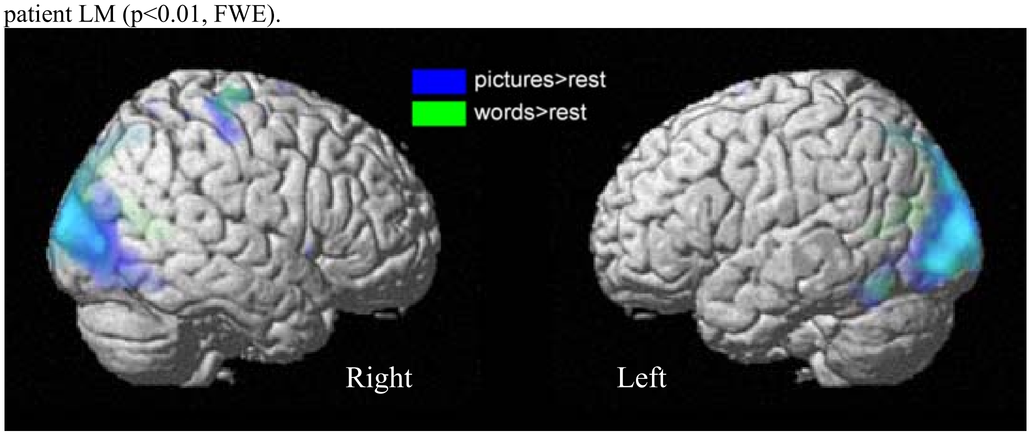
Figure 6.13. Sagittal, coronal and axial slices showing activation for words (words>rest; top) and pictures (pictures>rest; bottom) for patient CB.

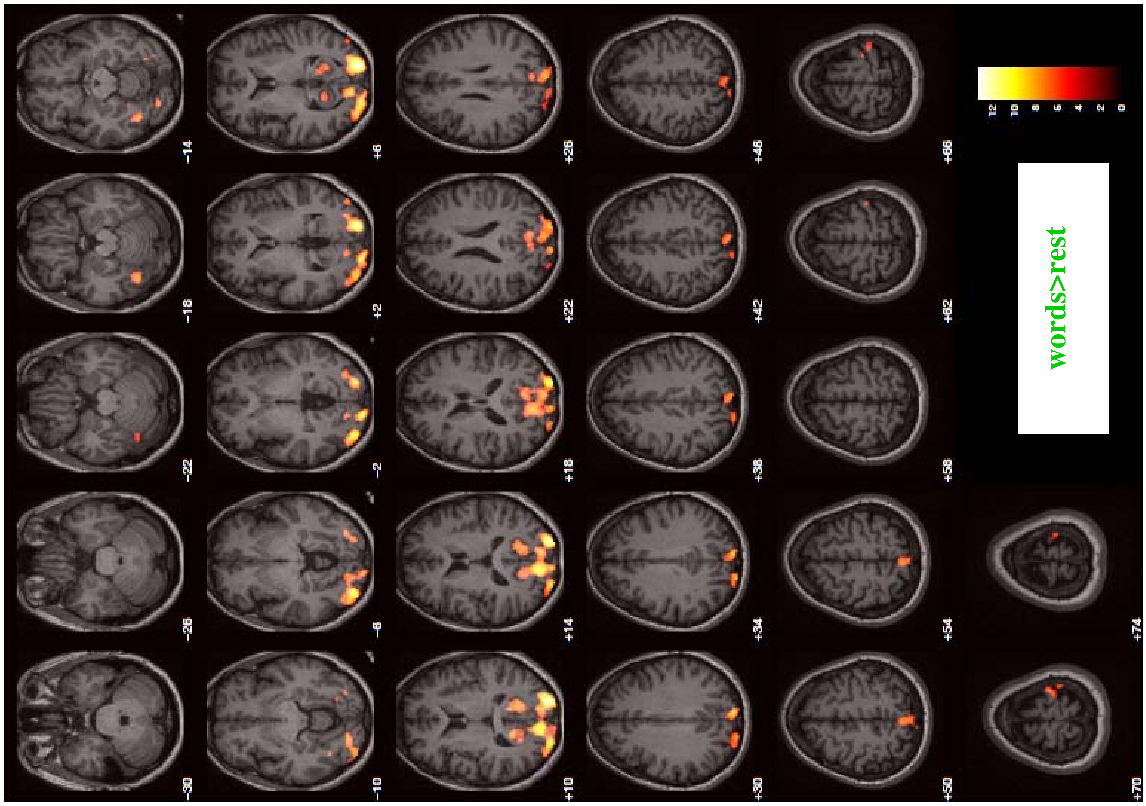


Patient **LM** had a small lesion in the insula and adjacent IFG. In the word task, patient LM showed activation in bilateral occipital regions, the left fusiform gyrus and the right post-central gyrus ($p<0.01$, FWE; figure 6.14-6.15). When comparing the word rhyme judgement to baseline, LM did not show any significant activation at the specified threshold ($p<0.01$, FWE). However, with a lower threshold ($p<0.001$, uncorrected), the right post-central activation was found and the occipital activation disappeared.

In the picture task, LM showed activation which was similar to the one obtained during the word rhyme judgement task. This included bilateral occipital and right post-central activation. A small locus of activation was also found in the right insula ($p<0.01$, FWE, figure 6.14-6.15). When comparing the picture rhyme judgement to its baseline condition, LM showed no significant activation at the specified threshold ($p<0.01$, FWE).

Figure 6.14. Brain activation during the word rhyming task (on this page, in green; on the next page, on the right), and during the picture rhyming task (on this page, in blue; on the next page, on the left) for patient LM ($p<0.01$, FWE).




words>rest

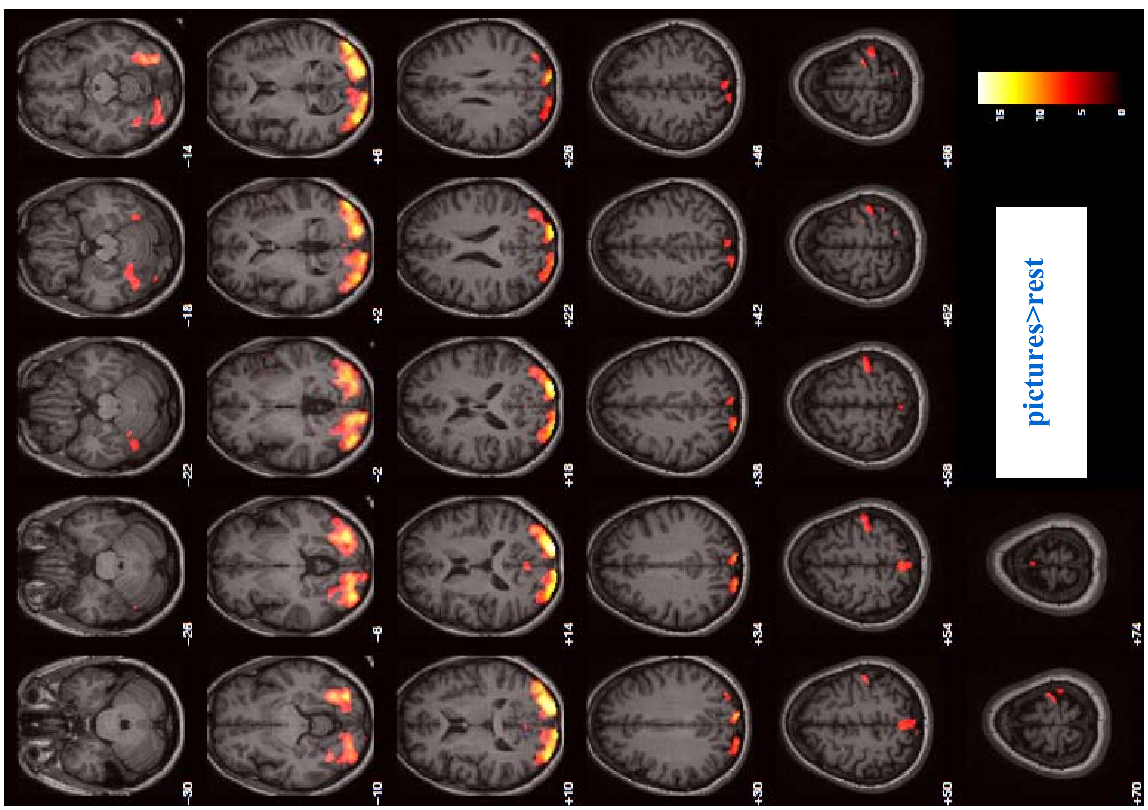
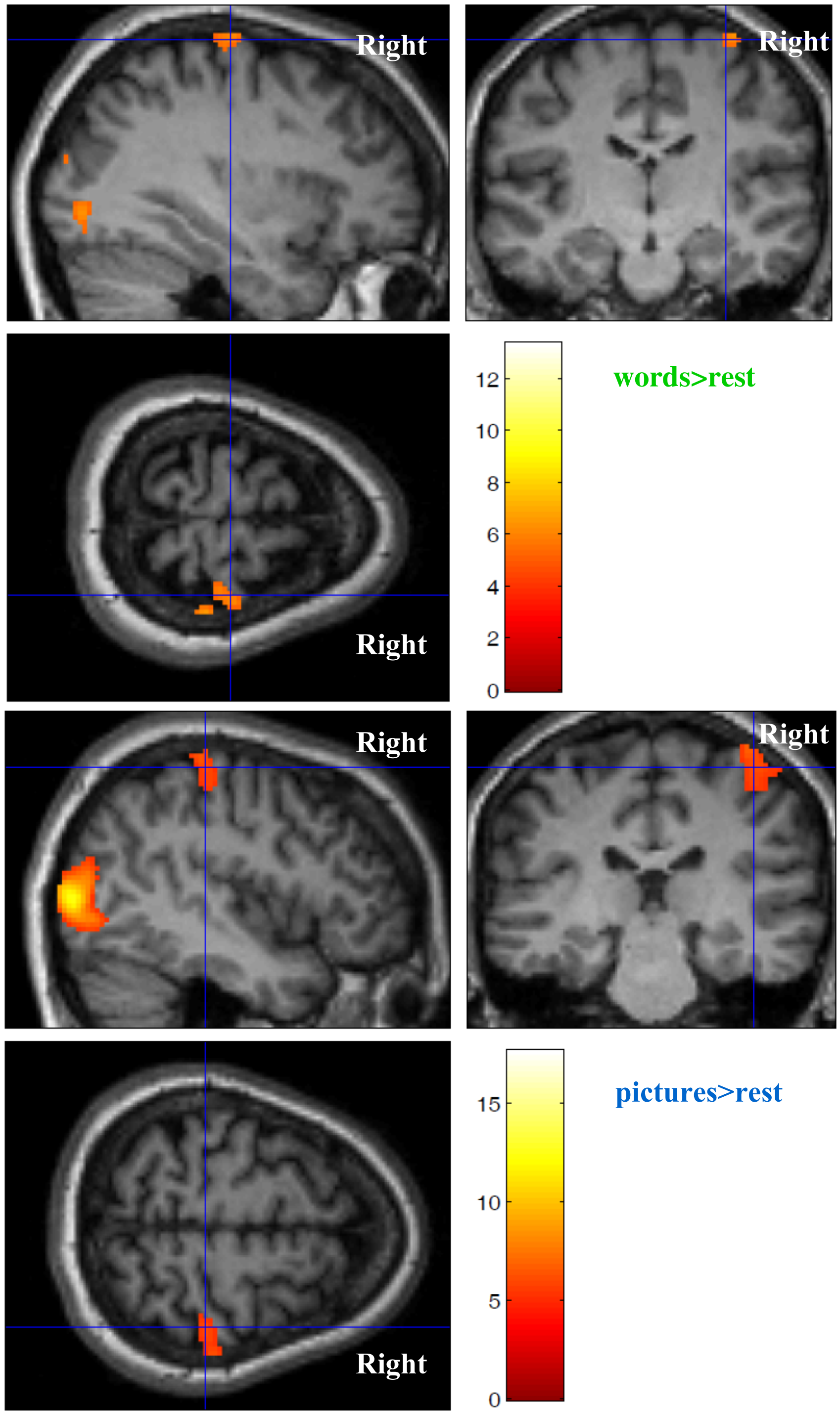

pictures>rest

Figure 6.15. Sagittal, coronal and axial slices showing activation for words (words>rest; top) and pictures (pictures>rest; bottom) for patient LM.



4. Discussion

This study explored the functional neural correlates of inner speech in healthy participants, ageing and stroke. In the next sections, the results from the various populations will be explored separately.

4.1 Neural Correlates of Inner Speech in Healthy Participants

The inner speech tasks were correlated with activation in three left hemispheric brain regions: the fusiform gyrus and inferior temporal lobe, inferior parietal cortex and the IFG. Activation in left inferior basal regions, and especially in the fusiform gyrus, was found in the word rhyming task. It is tempting to interpret this activation as related to the reading circuitry, demonstrating the involvement of the controversially named Visual Word Form Area (VWFA) in reading. However, similar activation was found in the picture rhyming task, and importantly, in the conjunction analysis. The interpretation of these findings therefore requires a broader framework. The left fusiform gyrus was associated with naming pictures in fMRI studies (Bookheimer, Zeffiro et al. 1995; Moore and Price 1999; Etard, Mellet et al. 2000), lesion studies (Raymer, Foundas et al. 1997; Foundas, Daniels et al. 1998), and studies using electrical stimulation (Luders, Lesser et al. 1986; Burnstine, Lesser et al. 1990; Luders, Lesser et al. 1991). Activation in this region was also found during a word rhyme judgement task (Booth, Burman et al. 2002). The tasks in these studies require phonological retrieval, or access to the phonological output store. The involvement of this area in phonological retrieval is further supported by a conjunction analysis study (Price and Friston 1997) and a meta-analysis of picture naming (Indefrey and Levelt 2004). An interesting study examined the effects of direct cortical stimulation of the left basal temporal cortex, on various language tasks in three Japanese patients (Usui, Ikeda et al. 2005). Relevant to the current discussion are the two Japanese writing systems; Kana and Kanji. Kana writing is phonological, similar to an alphabet, and can therefore be used for writing both words and non-words. Kanji characters, on the other hand, are strongly associated with semantics, and they are read differently depending on the context. Following a left basal temporal stimulation, patients were unable to name pictures and were partially impaired in reading aloud. Electrodes which induced complete impairment in reading aloud Kanji but only partial impairment in reading aloud Kana also induced impairment in matching Kana characters to pictures but not in matching Kanji characters to pictures. These results give further support to the idea that inferior and basal temporal regions are involved in phonological retrieval. Reading aloud requires phonological retrieval. Kana characters give information about the phonological form of the word which supports the retrieval process. Such support is not available with Kanji characters. As a result, complete impairment in Kanji reading, and only partial impairment in Kana reading, was observed. Furthermore, in the process of word-picture matching, Kanji requires no phonological access. A direct mapping of character to semantics can take place, allowing a successful character-picture matching. In Kana, on the other hand, mapping of orthography-to-phonology, and of phonology-to-semantics need to take place. The former requires phonological retrieval which was disrupted by the electric stimulation, therefore causing impairment in the Kana picture matching task (Usui, Ikeda et al. 2005). In summary, the study of Japanese reading together with other previous studies presented above, implicate the importance of the inferior temporal region in phonological retrieval.

One might wonder why this region was not highlighted in the structural studies presented in chapters 4 and 5. The answer might be a simple anatomical one: The fusiform gyrus is relatively protected from ischemic damage since it receives blood supply from both the posterior and middle cerebral arteries. Most patients participating in this study did not have a lesion in this area. This meant that the region was not analysed in the VLSM study and was also unlikely to be highlighted in the VBM study. Therefore, it might still be necessary for the task, and without examining patients with selective impairment in this region it is impossible to determine whether the area is necessary or contributing to task performance.

The second area of activation was found in inferio-parietal regions, including the supramarginal gyrus and the angular gyrus. The putative role of these regions in the process of inner speech production was discussed in previous chapters. In short, it was suggested that the supramarginal gyrus is responsible for binding speech perception and speech production, therefore acting as a link between the phonological input and output stores - a link which is essential for the production of inner speech, as described in chapter 3. In a recent review of 100 fMRI studies, the supramarginal gyrus (bordering the posterior planum temporale) was suggested to be involved in subvocal articulation (i.e. inner speech). Furthermore, it has been shown that activation is enhanced when speech production is more demanding. This suggests that the region is involved in auditory monitoring (overt or inner; Price 2010). Furthermore, it was shown that this activation is not dependent on working memory (Buchsbaum and D'Esposito 2009).

Activation of the angular gyrus is interpreted as providing semantic support during lexical access. Recent reviews of fMRI studies support this idea in suggesting that the angular gyrus plays a role in top down semantic activation, supporting speech production and comprehension (Binder, Desai et al. 2009; Price 2010). The fMRI paradigm in this study included only real words, and it is therefore impossible to determine, based on the generated data, whether the angular gyrus activation is specific to words. However, previous studies of non-word rhyme judgement did not find angular gyrus activation (Pugh, Shaywitz et al. 1996; Poldrack, Temple et al. 2001; Owen, Borowsky et al. 2004). Moreover, many studies that compared words and non-words directly, found significantly greater activation in the angular gyrus during word processing (reviewed in Price 2010). Lastly, angular gyrus activation was found when comparing sentences with plausible versus implausible meanings (Mashal, Faust et al. 2009) and when comparing auditory sentences to unintelligible speech (Obleser and Kotz 2010). Together, these studies suggest that the angular gyrus has a semantic role. Non-words were not used in this study because many patients can read words more easily than they can read non-words. Moreover, it was required that the same set of words be used in the pictures and written words conditions. Naturally, non-words cannot be presented in pictures. Lastly, patients are more likely to perform the task successfully if the stimuli are presented to them before entering the scanner. It is easier to practice the production of real words than non-words, with the latter more likely to be forgotten. Taken together, the limitations of the subject population and the materials prevented the use of non-words in this study.

As expected, activation was also found in the left IFG. This included the pars opercularis (BA44) and pars triangularis (BA45). It has been suggested in previous chapters that the pars opercularis is involved in phonological processing of speech production, especially when conscious inner speech is required. Price (2010) distinguishes between the dorsal pars opercularis, which is responsible for event sequencing during speech production and comprehension, and its ventral part, which is involved in planning articulation. However, she admits that the distinction is somewhat arbitrary. In this study, activation in the pars opercularis during the word rhyme judgement task was seen in both the ventral and the dorsal parts. During the picture rhyme judgement, activation in the dorsal part was found for all participants while activation in the ventral part was found only for the older controls. Activation in the pars triangularis was found in both tasks. Studies suggest that the pars triangularis has a role in semantic retrieval (Poldrack, Wagner et al. 1999; Wagner, Koutstaal et al. 2000; Wagner, Pare-Blagoev et al. 2001; Devlin, Matthews et al. 2003), although others suggest that semantic processing only takes place in more anterior parts, encompassing BA47 (Binder, Desai et al. 2009). The role of the pars triangularis in inner speech production might be similar to that of the angular gyrus: Both provide top-down semantic support. However, they might play a role in different stages of the process. While the IFG is involved in word retrieval (Poldrack, Wagner et al. 1999; Wagner, Koutstaal et al. 2000; Wagner, Pare-Blagoev et al. 2001; Devlin, Matthews et al. 2003), the angular gyrus is more involved in comprehension (Binder, Desai et al. 2009). With regard to inner speech, it is suggested that IFG activation takes place prior to angular gyrus activation, with the former modulating processes in the output phonological store and the latter sending top-down information to the input phonological store.

The pars triangularis of the left IFG was not highlighted in the structural studies presented in chapters 4 and 5. As discussed above, the main difference between lesion and functional studies is that while the former identifies areas which are necessary for a specific cognitive function, the latter reveals which areas are contributing to the performance of the task. In light of this distinction, it is suggested that the pars triangularis of the left IFG is contributing, but is not necessary, to inner speech production. A TMS study supports this suggestion. TMS to the anterior left IFG influenced the response time, but not the error rate of responses to a semantic task (Devlin, Matthews et al. 2003). Furthermore, patient with left IFG lesions do not typically show substantial semantic impairments (Benson and Ardila 1996; Price, Mummery et al. 1999). Taken together, it is concluded that the anterior left IFG contributes to inner speech processing: it provides semantic support at the level of word retrieval.

Lastly, extra-sylvian activation was found in the MFG, especially in the area of the DLPFC and BA6. Right hemispheric activation was mainly found in the older control group and will be discussed separately below, when considering ageing effects.

These results largely replicate the results obtained in the VLSM and VBM studies. The three approaches differ to a large extent in their methodology, strengths and biases. In this work, the studies also differed to some extent in the participants who took part in each study, in statistical power and in the materials used. Despite these differences, the results were largely replicated, a finding which supports their reliability.

4.2 Neural Correlates of Inner Speech in Ageing

The effects of ageing were explored by comparing the activation patterns of two groups of participants: younger adults (mean age: 25) and older adults (mean age: 64). Behaviourally, the groups showed no significant differences. Their RT and error rates were similar in all four tasks (two rhyme judgement tasks and two baseline tasks). While the young participants did not show any activation which was significantly greater than that of the older controls, the older controls showed many areas of activation which were significantly greater both in magnitude and extent, when compared with the younger participants. These ageing related activations can be divided into three types: 1.) Frontal lobe activation (bilateral DLPFC in the word rhyme judgement; left BA6 in the picture rhyme judgement). 2.) Right hemispheric language homologue activation (right insula and right pars opercularis in the word rhyme judgement; right pars triangularis and opercularis in the picture rhyme judgement), and 3.) 'Other' activation (midline and left cerebellum in the word rhyme judgement; bilateral cerebellum, left pre-central gyrus, right pre- and post-central gyrus in the picture rhyme judgement). These results are discussed within the framework of ageing theories which were presented in the introduction of this chapter.

By and large, ageing-related increases in the level or extent of activation, together with normal behaviour, are usually interpreted as compensatory mechanisms. Accordingly, the increased activation in the older participants, seen in this study, compensates for potentially impaired performance. It then allows the performance to be maintained at equal level to those of the young participants.

However, the correlation analysis reveals a more complicated picture. In the picture rhyming task, a reduction in grey matter density in the right cerebellum correlated with increased right IFG activation in the pars triangularis (BA45), but only in the older participants. In the absence of any behavioural correlate, this result is difficult to interpret. However, it suggests that cerebellar grey matter loss may have impaired performance and this in turn was related to compensatory activation in the right inferior frontal cortex. Further support for this interpretation comes from evidence about the role of the right cerebellum in language processing. The right cerebellum is known to be involved in motor control of speech (Palmer, Rosen et al. 2001; Wise 2003; Foki, Gartus et al. 2008), but many studies also show its involvement in non-motor language processing (Petersen, Fox et al. 1989; Papathanassiou, Etard et al. 2000; Price 2000; Palmer, Rosen et al. 2001; Shuster and Lemieux 2005; Frings, Dimitrova et al. 2006). It has been suggested that the role of the cerebellum in language is to process internal representations (Strick, Dum et al. 2009), which is especially relevant in this study. Others have demonstrated right cerebellar involvement in the processing of grammar (Silveri, Leggio et al. 1994; Marien, Saerens et al. 1996; Zettin, Cappa et al. 1997; Marien, Engelborghs et al. 2000; Justus 2004). The right cerebellum is therefore often seen as part of the mostly left (cortical) hemispheric language network. Language impairments following cerebellar infarct are often explained as resulting from *crossed cerebello-cerebral diaschisis* (Marien, Saerens et al. 1996; Marien, Engelborghs et al. 2000; Baillieux, De Smet et al. 2008), in which a cerebellar infarct impairs the normal activity in the left frontal regions. With reference to the results of this study, it is suggested that a reduction in right cerebellar grey matter density induces alterations of normal activity in the left inferior frontal cortex, which in turn is compensated by enhanced activation in the right inferior frontal cortex. This allows performance to be maintained at normal levels. These results support both the *Hemispheric Asymmetry Reduction in Older Adults* (HAROLD) model (Cabeza 2002) and theories emphasising the importance of frontal activation in older participants (West 1996).

However, in the word rhyming task, increased activation in the right inferior frontal cortex (pars orbitalis; BA47) was correlated with slower response time, only in the older participants. This suggests that in this case, right IFG activation is a marker of processing inefficiency, affecting speed of response, but not its accuracy.

Grady (2008) suggests that changes in activation patterns may reflect the use of different strategies. However, she does not point out what these alternative strategies are. Moreover, there is no clear theoretical reason to believe that the way people process language changes with age. One could even argue that patterns of cognitive processing become more rigid with age. In this study there was an attempt to control for possible strategies that subjects might use when performing the task.

Firstly, participants were asked to determine whether two words rhyme by producing inner speech. An alternative strategy might be to compare the orthography of the two words (and in the case of pictures, to retrieve the orthography and compare the retrieved output). In order to prevent participants from using this strategy, the word pairs had an orthographically dissimilar ending. Participants were also asked not to use articulation as an aid in performing the task. To ensure this, all participants practiced the task with tens of word pairs until it was confirmed that they were not using articulation. Moreover, they were asked to actively hold their mouth closed while in the scanner (although no bite-bar or similar apparatus was used). Other potential strategies and the way they were dealt with are discussed in the methods section of this chapter. In summary, the study was designed in a way that would reduce the use of alternative strategies as much as possible.

Ageing effects were also found in areas which are not language related or language homologues. Here it was shown that left BA6 activation during the word rhyming task was correlated with increase in response time for the older, but not for the young, participants. This finding is similar to the correlation described above, between the right IFG activation and response time. Again, such correlation suggests that the activation reflects inefficiency in neuronal processing. During inner speech production there is an absence of overt auditory or sensory feedback. It was proposed that this absence of feedback increases the reliance on ‘feed forward’ mechanisms, and that the premotor cortex might be involved in providing such ‘feed forward’ support (see discussion in chapter 5). However, the correlation with response time suggests that this recruitment is not entirely efficient, and, although it might support response accuracy, it does not support response speed.

Activation in other brain regions might reflect similar processes which are not essential to inner speech production but support the process. For example, pre- and post-central activation might also be related to sensory and motor feed forward mechanisms. However, these would have to be studied separately.

Lastly, it was found that ageing effects were greater and more widespread in the picture task, compared to the word task. This result confirms the hypothesis that tasks using pictures are more likely to show ageing effects than those using words, since the main ageing-related difficulty in speech production is at the level of word retrieval, while reading per se remains unimpaired even in very old age.

In summary, a correlation between increased right hemispheric activation in a language homologue region and decrease in grey matter density in the language network, suggests that this age-related activation is beneficial for performance. On the other hand, a correlation between increased activation in the right inferior frontal cortex and slowing of response time suggests that this activation reflects processing inefficiency. These results are consistent with Wierenga et al.'s (2008) fMRI study of word retrieval, which demonstrated that ageing related activation can reflect both inefficiency and compensation. Their study demonstrated that activations in the same subject and in the same task, but in different brain regions, can reflect these two different age-related processes. The study presented here expands on their results by showing that activations in the same subject and in a similar brain region can reflect the two opposite processes, depending on the task.

4.3 Neural Correlates of Inner Speech in Patients with Aphasia

Imaging studies of stroke patients often discuss results within similar conceptual frameworks as the ones used to discuss ageing effects. The terms: compensation, processing inefficiency, lateralisation and alternative strategy, among others, are used to interpret results from both populations. In this study, patients with aphasia demonstrated peri-lesional activation, contra-lesional activation, and normal activation (comparable to that of the respective control group). Each of these effects will be now discussed in turn.

Peri-lesional left activation was most pronounced for patients IB, AD and BA. All three patients had perisylvian lesions. Patients IB and AD had lesions affecting the left IFG, among other cortical structures. Patient BA had a lesion affecting the left inferior parietal structures. This means that all three patients had a lesion in at least some of the areas found to be essential for inner speech production. The patients' error rate was close to zero and it is therefore suggested that the peri-lesional activation is beneficial in these cases. It has long been known that peri-lesional activation has a positive effect on recovery of function (Furlan, Marchal et al. 1996; Iglesias, Marchal et al. 1996; Nudo, Wise et al. 1996; Price, Mummery et al. 1999; Warburton, Price et al. 1999; Saur, Lange et al. 2006; Meinzer, Flaisch et al. 2008). The current findings replicate the results of Perani and colleagues (2003) who showed the existence of peri-lesional activation during the performance of a non-word rhyme judgement in a recovered patient with aphasia. In addition to peri-lesional activation, patient IB showed extensive activation across the entire cortex in all tasks. Although the patients as a group did not differ in their performance (both RT and error rate) from the control group, IB's response times were more than 3 standard deviations above the overall controls' mean in both rhyming tasks. This slowing of response was also evident during the behavioural testing: while IB performed at ceiling in many tasks, his responses were notably slow. It is suggested that the vast activation seen in his cortex reflects a more effortful response. Based on the imaging results, it is impossible to determine in which phase of processing the slowness occurs, and it is possible that some levels of processing are more affected than others or that the slower responses reflect global slowing. The 'cognitive slowing' theory (Salthouse 2003) suggests that cognitive processes slow down with age, a theory which was supported by an ERP study of word retrieval (Galdo-Alvarez, Lindin et al. 2009). This process might be enhanced by stroke, and might be evident in both response time and overall brain activation.

Contra-lesional activation was found for patients IB, BA and CB. While for patient IB this activation was non-specific and may be attributed to processes such as increased attention and effort, as discussed above, for patients BA and CB the right hemispheric activation was more localised. Patient BA had a left inferior parietal stroke which affected the angular and supramarginal gyri. His right hemispheric activation mirrored the lesion. His performance was at ceiling, with response times that were faster than the overall controls' mean. It should be noted that BA is also the only patient who showed complete recovery with no impairments in any of the behavioural tests. Taken together, it is suggested that this right hemispheric activation is either beneficial or epiphenomenal and this could be established using a method such as TMS. Patient CB had a subcortical stroke affecting mainly the left putamen and adjacent white matter. This is surprising considering his initial presentation of severe global aphasia. It is proposed that some of his symptoms might be explained by damage to white matter tracts. The idea that damage to brain connections can cause aphasia is not new. It was first introduced by Wernicke (1874) and was further developed by Lichtheim (for a review see Catani and Mesulam 2008). Most studies which relate structure to function focus on the grey matter. However, lesions to the grey matter cannot always explain the complex symptoms presented by some patients. For example, Mao et al. (2007) found that a patient with aphasia had no signs of damage on regular MRI sequences (T1 and T2), six months after a car accident that caused his brain injury. However, performance on language tasks was clearly impaired, especially in the domain of language production (verbal fluency and naming). DTI analysis revealed a reduction of fractional anisotropy (FA) values in the left frontal region, especially around the areas that were activated during performance of a language task. Similarly, patient CB's symptoms might be a result of damage to white matter tracts which affect the normal functioning of some perisylvian regions. This in turn can result in the development of contra-lesional activation, as seen in this study. A DTI analysis may help in further understanding and interpreting the relation between the lesion, symptoms, and activation for patient CB.

Contra-lesional activation was found in the two studies of inner speech in post-stroke aphasia reported above and both interpret this activation as being beneficial for the patient's recovery (Calvert, Brammer et al. 2000; Perani, Cappa et al. 2003). However, the role of contra-lesional activation has long been debated, a debate which has been briefly mentioned in the introduction to this thesis. While some argue that contra-lesional activation is beneficial for performance (Cappa, Perani et al. 1997; Calvert, Brammer et al. 2000; Perani, Cappa et al. 2003; Winhuisen, Thiel et al. 2005), other suggest that it is a maladaptive reaction (Naeser, Martin et al. 2004; Price and Crinion 2005) and can even be a predictor of poor recovery (Kurland, Naeser et al. 2004). In a study of aphasia therapy it has been shown that both correct responses and therapy-related effects correlated with contra-lesional activation, suggesting that this activation is beneficial for recovery (Meinzer, Flaisch et al. 2006). Applying TMS to the right IFG of recovered stroke patients impaired performance for some patients (Naeser, Martin et al. 2005; Winhuisen, Thiel et al. 2005). In a cohort of patients with left hemispheric brain tumours, applying TMS to the right hemisphere affected language performance only in patients with long disease duration (Thiel, Habedank et al. 2006). This finding is consistent with current models which argue that contra-lesional activation takes time to develop and therefore occurs only in sub-acute and chronic post-stroke stages (Marsh and Hillis 2006; Saur, Lange et al. 2006). The TMS study of tumour patients suggests that the slowly developing deficit enabled the right hemisphere to gradually take over some of the functions of the left hemisphere. When the disease developed quickly, this shift did not occur (Thiel, Habedank et al. 2006). Lastly, some studies have shown that after recovering from a first stroke, a second stroke affecting the other hemisphere can cause re-emergence of the initial symptoms (Levine and Mohr 1979; Lee, Nakada et al. 1984; Basso, Gardelli et al. 1989).

Although these studies suggest that the contra-lesional hemisphere initially took over the functions of the lesioned hemisphere, more recent studies have questioned this interpretation, suggesting that detailed examination of the site of lesion or pre-stroke wiring can explain these rare results (Yamamoto, Takasawa et al. 2007). In summary, there is evidence that contra-lesional activation may be beneficial for recovery of language function in some cases, although this depends on the recovery stage and size and site of the initial damage, among other factors.

Lastly, it is widely agreed that patients with small lesions, or lesions in areas which are less crucial for language function, show better recovery (Heiss and Thiel 2006; Marsh and Hillis 2006). Such patients typically show activation patterns similar to those seen in healthy participants. In this study, such activation has been demonstrated by LM, a 48-year-old woman with a small lesion in the region of the left IFG and insula. Activation patterns in the two rhyming tasks were largely similar to that of the younger control group. It should be noted that on examination of individual results it was found that some of the young controls showed mainly occipital activation, similar to patient LM. This was not the case for the older controls who all showed perisylvian activation as well as occipital activation (individual data not shown). Hence, LM, a young patient with a small lesion and good recovery of language function, showed activation which was comparable to the young control group.

In summary, the patients who participated in this study showed various patterns of activation as response to the language task. These included peri-lesional, contra-lesional and normal activation. The various patterns of activations can be explained by the patient's behavioural performance and lesion location.

4.4 Main Study Caveats

On many occasions, the comparison of the main tasks to their baselines resulted in no significant activation at the predetermined thresholds. This is despite the fact that the baseline task had no explicit linguistic aspect. It was designed to control for effects which were of no interest to the study such as visual and motor activation. Such baseline was expected to increase the sensitivity of the study, in comparison to a baseline such as rest (Price, Devlin et al. 2005) but this was clearly not the case. It is suggested that the failure of the baseline condition in controlling only for effects of no interest, lies in the time window given to each trial. In the study, each trial was given a time window of 7.3 seconds. Participants responded to trials in the baseline tasks in less than 1.5 seconds, on average. This left participants with more than 5.5 seconds in which they were not required to perform any specific task, a time window which can be considered substantial in terms of both fMRI and behavioural studies. In this time window participants could engage in uncontrolled cognitive processes. This will inevitably create noise which potentially can mask activation when the baseline task is used in the analysis. It should be noted that this time window was determined based on preliminary work done with aphasia patients. However, patients who later participated in the fMRI experiment performed the task much faster, and in fact, did not differ in their response time from the control group. It is therefore concluded that future studies should use shorter time windows.

4.5 Conclusions and Future Directions

This study demonstrated that ageing brings about various changes in brain activation which can largely be classified as: 1.) An increase activation in left hemispheric language areas, 2.) Right hemispheric activation in language-homologue regions and, 3.) Activation in other areas which are not activated by young participants. These patterns were paralleled by the stroke patients, who showed 1.) Peri-lesional (left hemispheric) activation, 2.) Contra-lesional (right hemispheric language homologue) activation, and, 3.) Activation outside these regions which might reflect increase in effort or attention, or the use of feed forward and feedback mechanisms to support inner speech production.

Single-case and group studies can both contribute to our understanding of brain plasticity in general, and language processing after stroke, in particular. However, there is enormous variability among patients in their lesion location and size, pre-morbid abilities, age, therapy given, and other factors which might be relevant to language recovery. Pooling many patients together in a group study can potentially obscure these differences. At the same time, a single-case study cannot reflect this variability within the patient population. Therefore, multiple-case studies can potentially tackle and explore this variability within the patient population. This study demonstrates this approach and highlights the importance of studying patients with different lesions and abilities, using the same paradigm and materials.

Chapter 7
General Discussion

Things should be made as simple as possible, but not simpler.
–Albert Einstein

1. Background

This thesis explored the behavioural manifestations and neural correlates of inner speech. The original motivation for this research project was clinical. Patients often complain that while they can hear their inner speech, they are sometimes unable to speak out loud. The opposite complaint is rarer, but still exists: some patient feel that the stroke made them 'lose the little voice in their head'. Since inner speech is fundamental for many language and other cognitive functions, this study's goal was to explore inner speech in post-stroke aphasia.

However, there were other motivations for this study. Pre-operative assessments of language in patients undergoing neurosurgery (for epilepsy or tumours) are often performed to define the boundaries of the language regions, in order to preserve these during surgery. These assessments use only overt language assessments while presuming that the boundaries of the inner speech system are the same.

In the area of neuroimaging, language is one of the more prolific areas of research, resulting in hundreds of studies published every year. However, there is a fundamental problem with many studies looking at speech production. Since producing overt speech in the MRI scanner is methodologically problematic, many studies use covert response, relying on the assumption that inner and overt speech are identical at every level except in articulation. However, this clearly needed to be evaluated in a systematic study.

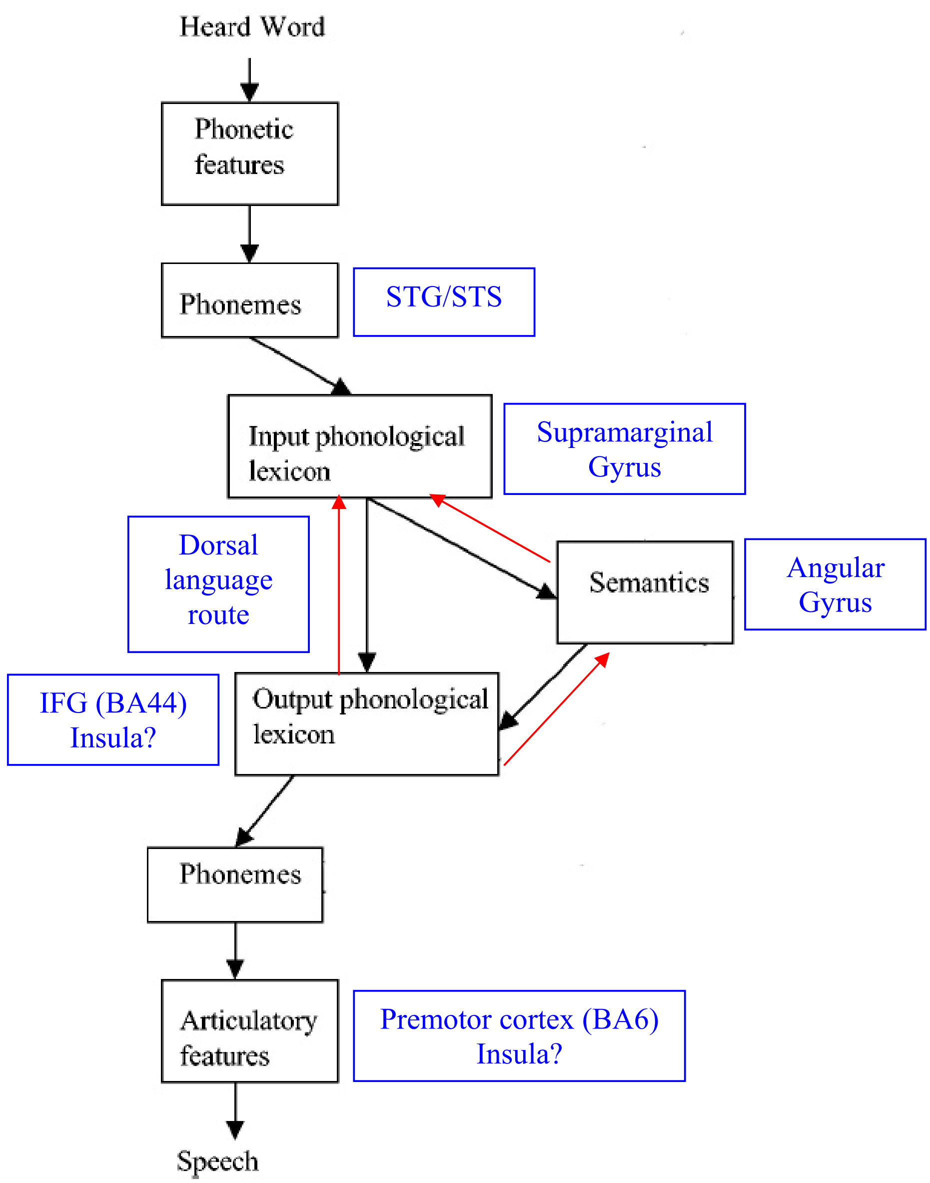
Lastly, cognitive models of language production and comprehension disagree about the nature of inner speech. Studying patients with aphasia can often contribute to debates about theoretical models. Moreover, imaging studies have the power to guide cognitive models. Although Vigliocco and Hartsuiker (2002) rightly pointed out that '…there is not necessarily a direct correspondence between cognitive systems and their neural implementation', they also argued that '…psychological models should be at least consistent with biological constraints' (ibid, p. 465). In this regard, imaging studies have the power to inform cognitive studies by providing these 'biological constraints'.

A literature review looking at behavioural and imaging studies of inner speech in healthy participants and patients with aphasia, identified a clear lack of consistent, well-controlled data. It was also noted that in the past, inner speech was a major topic for research. The interest in inner speech, however, seemed to wane at the turn of the last century, with fewer and fewer neurologists, aphasiologists and cognitive psychologists dedicating their research to the topic. Behaviourism probably brought about a further decline in the interest in inner speech, a cognitive function which might be considered a *'black box'*. Marcel Kinsbourne, in the Millennium issue of *Brain and Language,* noted that "…inner speech is neglected at Millennium's end, not even mentioned in the Handbook of Neurolinguistics (1998). The dying century has witnessed an inward gradient of research focus, from overt behavior, through the intellectual models of the cognitive revolution, to a fin de siècle finale of swelling interest in consciousness. Extrapolating to the next millenium of research, I predict a shift of focus from the mechanics and organisation of utterance to its more elusive origins in thought. Investigators will address the functions of inner speech, and its precursors in ontogeny, speech for self. What does inner speech contribute beyond flexible and economical recoding, and how does disabled inner speech harm cognition?" (Kinsbourne 2000, p.120). Therefore, this thesis aimed to contribute to the two spheres of research: cognitive investigations and brain imaging of inner speech.

2. Main Findings

This study explored the behavioural and neural manifestations of inner speech. It was aimed at contributing to current anatomical and cognitive models of inner speech. The results are presented below with their relation to current models of language processing. The results are presented within the framework of the model presented in chapter 1 (figure 1.6). This model is by no means the only, nor the most cited model in the area of language processing. However, it encompasses many influential models which are widely used today, both in the literature and in speech and language therapy. The model is a modification of the influential model by Ellis and Young (1988), and is adapted from Martin (2003). Here I focus on auditory language processing and speech production, the processes most related to inner speech production, over and above lead-in processes. Marked in red are proposed connections based on the data in this study. In blue are brain regions that are suggested to play a role in the processing of inner speech. The executive control system, which is located in the DLPFC and is responsible for conflict resolution and attention modulation among other functions, is not included in this figure.

Figure 7.1. Adaptation of the language processing model to results from this study.



Based on the behavioural work presented in chapter 3, it was suggested that conscious inner speech is processed by transferring information from the output phonological store to the input phonological store. It was also argued that models which do not differentiate between the two, such as some connectionist models, cannot easily account for the data. It was also suggested that the semantic system has a two-way connection with the phonological stores. This is similar to models developed by Dell (Dell 1986; Dell and Oseaghdha 1992) or Seidenberg and McClelland (Seidenberg and McClelland 1989), in which feedback connections allow top down and bottom up interactions between the components of the system. Therefore, the revised model suggested here, incorporates features which can be perceived as 'connectionist' and some which can be conceived as 'non-connectionist'.

The imaging results were notably replicated in the three studies (VLSM, VBM and fMRI), albeit the use of different methods, different materials and somewhat different participants, in each study. The suggested roles of the various brain regions are not novel. Previous studies have demonstrated the involvement of the STS/STG in phoneme discrimination (reviewed in Price 2000; Boatman 2004; Demonet, Thierry et al. 2005), the angular gyrus in semantics (Price 2000), BA44 in phonological processing in the context of speech production (Mummery, Patterson et al. 1996; Huang, Carr et al. 2002; McDermott, Petersen et al. 2003; Sato, Baciu et al. 2004), BA6 in motor planning (Solodkin, Hlustik et al. 2004; Sharma, Simmons et al. 2009), the insula in planning articulation (Dronkers 1996; Wise, Greene et al. 1999; Price 2000; although see also Hillis, Work et al. 2004) and phonological processing (Price and Friston 1997; Price 2000; Price, Devlin et al. 2005; Vigneau, Beaucousin et al. 2006; Shafto, Burke et al. 2007) and the supramarginal gyrus in input phonology (Xiang, Fonteijn et al. 2010). Thus, the study does not suggest a completely new role to any specific brain region. Its importance is in providing a new understanding of the way in which the various components of the language processing network work in concert to allow and support inner speech processing.

3. Implications of the Current Research

This research project has implications for clinical practice, imaging studies, and cognitive models of language processing. Each of these domains is discussed separately below.

It is generally assumed, that when patients with aphasia make overt speech errors this is due to overall impaired language abilities. However, this study showed that inner speech can remain intact even when there is a marked deficit in overt speech. It is suggested that inner speech be evaluated routinely both for patients with aphasia and for patient who are about to undergo neurosurgery for brain tumours or epilepsy. This practice is not common among speech and language therapists and neurosurgeons. Such assessment can potentially influence the diagnosis, prognosis and even therapy plan given to patients with aphasia. Imaging studies also have a promise in the field of aphasia diagnosis, prognosis and therapy. Today, our ability to give accurate prognosis and to fit each patient with the most beneficial therapy regime are, both, highly limited. However, many researchers are already investigating the option of using brain imaging as a tool to increase reliability of diagnosis and prognosis and to guide therapy. Using the information acquired through many carefully designed imaging studies, one will be able to incorporate the knowledge about the patient's lesion site, size, and brain activation, among other factors, into the overall picture of the patient's language deficits and potential for rehabilitation. It has been suggested that developing a multi-centre data base of patients data will be the most powerful way to contribute to such an endeavour (Price, Seghier et al. 2010).

The use of imaging paradigms as preoperative evaluations for tumour and epileptic patients due to undergo brain surgery is expanding (see Koppel and Buchel 2005, for a review). Conducting such an evaluation using fMRI has a few advantages. First, unlike the Wada test (Wada and Rasmussen 1960) and electrocortical stimulation, it is non-invasive and can be repeated many times. Secondly, the Wada test results in neurological complications in 1.3% of the cases. Furthermore, in 0.5% of all cases the damages are permanent (Willinsky, Taylor et al. 2003). Patient cooperation, psychological strain for the patient and extended craniotomies are the most common problems related to electrocortical stimulation (Foki, Gartus et al. 2008). fMRI has the potential to be an effective and convenient preoperative assessment tool, replacing the more problematic Wada test and electrostimulation. However, most fMRI paradigms employ inner speech tasks. In accordance with Foki et al. (2008) it is argued that in order to be able to use fMRI as an efficient and safe preoperative tool, it is important to understand the anatomical differences between inner and overt speech.

The results of this study may influence the construction of future language imaging paradigms for fMRI. Huang et al. (2002) stated that ‘it is incorrect to view the neural substrates of silent and overt speech as the same up the execution of motor movements, and it is therefore inappropriate to use silent speech as a motion-free substitute for overt speech in studies of language production’ (ibid, p. 50). The results of the research presented in this thesis, further support this statement and therefore have implications for imaging studies: it is suggested that inner and overt speech tasks have differences as well as overlaps in their brain localisation. Therefore, using inner speech tasks to study overt speech production may well result in misleading conclusions.

When designing imaging studies one should also take into consideration the effects of ageing. The results of this study clearly show that young and old participants differ significantly in the way their brain process language. However, often studies of stroke patients use young participants as their control group or discuss their results in comparison to studies of young participants. This is despite the fact that most stroke patients are above the age of 65. It is therefore suggested that studies of stroke patients use an age-matched control group and discuss results with relation to studies of ageing rather than studies of healthy young volunteers.

Lastly, this research has some potential implications for the area of cognitive models of language. It is suggested that models which define inner speech as dependent on the language production system alone, cannot easily account for the results of the behavioural and imaging studies presented here. The main suggested modifications to current language models are highlighted in the figure above (see figure 7.1).

4. Main Limitations of the Current Research

The main limitation of the studies presented in this thesis is the lack of sensitivity and specificity of some of the tests used. Specifically, working memory should be tested more thoroughly, ideally with tasks which do not require repetition. These should be applied to the healthy participants as well, since such data could have contributed to both the behavioural and the VBM studies. Ideally, such tests should be designed in such a way that will not create a 'ceiling effect' in the controls, nor a 'floor effect' in the patient population. It might be difficult, if not impossible, to find one such test. A battery of tests looking at working memory should be employed in future studies in this area. A second test with limited sensitivity in this study is the Apraxia Battery for Adults. While this test is useful in clinical settings, it seems not to be sensitive enough to small differences in motor speech abilities which might be relevant for this type of research. Again, a battery of tests for apraxia would be ideal to increase sensitivity in this domain.

5. Future Research

Three lines of research are suggested for future work: expanding the behavioural understanding of inner speech; using new imaging techniques to further our understanding of the neural correlates of inner speech; and, exploring the use of inner speech in speech therapy.

5.1 Behavioural Research on Inner Speech

In order to expand on the information gathered about inner speech in this research project, various levels and manifestations of inner speech should be explored. For example, one can explore different 'lead in processes'. In this study mainly written material was used, but future studies can, for example, also use pictures (as the ones used in the fMRI study). Moreover, there are other tasks which employ and examine inner speech, and using a vast battery of inner speech tasks has some major advantages. Firstly, a battery of tests can more easily isolate various cognitive components which are part of inner speech processing. Secondly, it will allow different patients to perform different tasks, ideally allowing wider inclusion of patients in the study. It can also assist in exploring the potential differences between conscious and unconscious inner speech. Lastly, employing various inner speech tasks can help in understanding how impairment in inner speech can influence reading, thinking, memory, self-referential activities and other cognitive functions which are highly dependent on inner speech. Tasks which can be used include the letter cancellation task, simpler homophone judgement tasks (for example, using a letter and its name: does 'I' and 'eye' sound the same? Does 'B' and 'bee' sound the same? or a pseudo-word and a real word: does 'brain' and 'brane' sound the same?). One can also ask participants to determine whether a single pseudo-word sounds like a real word (does the nonsense word 'brane' sound like a real word?), or whether a specific phoneme exist in a given word (can the sound 'f' be found in the word 'laugh'?).

Some preliminary work was done to develop a task in which only automatic speech is used. In this task, patients are asked to recite the days of the week according to a set pace. When stopped, they are asked to report which day they have arrived at. The motivation for developing this task came from the well documented finding regarding the behavioural dissociation between automatic and voluntary speech production in aphasic patients. Patients with impaired production of speech usually produce better verbal output when asked to present over-learned, automatic speech such as counting or reciting the days of the week (Benson and Ardila 1996), singing (Yamadori, Osumi et al. 1977; Hebert, Racette et al. 2003) and even praying (Hebert, Racette et al. 2003). As early as 1879, Hughlings-Jackson noticed that patients who could not produce voluntary speech, can sometimes still sing, curse, recite or produce highly stereotypical speech (e.g. saying 'goodbye' when a person leaves) (Hughlings Jackson 1879). Lum and Ellis (1994; 1999) found that a group of aphasic patients showed a consistent advantage in performance of automatic (non-propositional) speech tasks compared to matched voluntary (propositional) speech. The use of automated speech can therefore provide us with a means to evaluate inner speech even in very impaired patients.

5.2 Imaging of Inner Speech

The studies presented in this thesis provided information about the structural and functional neural correlates of inner speech. One of the main ideas suggested here is that inner speech is supported by the dorsal route for language. However, it was also noted that without a technique dedicated to examining white matter tracts, such postulation cannot be definitively made. Diffusion Tensor Imaging (DTI) is used to study connections between brain regions in the living human brain (Shimony, Snyder et al. 2004). DTI cannot determine whether a specific tract is afferent or efferent (Mori and van Zijl 2002) - a question specifically relevant for this study. However, at present, DTI is the only method to evaluate tract anatomy and tract changes in the living human brain.

DTI can be used as a tool in group-comparisons (Kim, Kim et al. 2006; Vernooij, Smits et al. 2007). It is therefore possible to examine whether connectivity values differ between patients with impaired inner speech and those with relatively preserved inner speech, and indeed from healthy controls. DTI can also be used as a tool for linking structure and function (Selnes, van Zijl et al. 2002). As discussed before, there is a long lasting debate regarding patterns of cortical activation in recovered and partially recovered chronic stroke patients. While some researchers argue that peri-lesional activation is necessary for function, others believe that remote and even contra-lesional activation is responsible for the recovered function. In this study, fMRI identified areas that are active in patients, but not in controls, during inner speech. A DTI analysis could examine whether any of these newly activated areas are connected to the spared language areas in the parieto-temporal junction and to the motor articulation areas in the pre-central gyrus and the insula. Peri-lesional areas have established pre-stroke connections to the language and motor areas. Therefore, recovery that is based only on these areas will not require the establishment of new connections or unmasking of existing but previously less active connections. However, if activation of remote intra-hemispheric and contra-hemispheric areas is necessary for recovery of behaviour then new connections must be activated. If the latter is found, it would imply that recovery after stroke involves plastic changes which make use of tracts connecting healthy cortical areas which are language related pre-stroke, to areas which 'take-over' the function of the damaged tissue. To examine this hypothesis, a combined functional-anatomical method which places a seed for white matter tracking in the area of fMRI activation, can be used (Staempfli, Reischauer et al. 2008).

Lastly, giving prognosis and studying recovery are today based mainly on information about damage to the grey matter. Catani and Mesulam (2008) suggest that DTI can further contribute to our understanding of recovery and to the ability to give accurate prognosis. For example, they suggest that greater symmetry in white matter tracts can lead to better recovery, with right hemispheric tracts compensating for loss of function in the left hemisphere.

5.3 Aphasia Therapy

As mentioned before, the main incentive for this research project was a clinical one. The study aimed to further enhance understanding of inner speech and the possible use of inner speech tasks in diagnosis and prognosis. However, inner speech can potentially be used in therapy as well. The incentive for looking into a new rehabilitation approach comes from the disadvantages of current therapies, discussed in the introduction. Using inner speech for language therapy will parallel promising developments that have already been made in research studying the rehabilitation of patients suffering from post-stroke motor deficits. While patients cannot produce a movement, they are sometimes able to imagine such a movement. Such 'motor imagery' is suggested to improve motor performance through the enhancement of brain activation (Jeannerod 1995; for a recent review see Sharma, Pomeroy et al. 2006). Drawing from this approach, it is suggested to test whether inner speech can enhance the activation in brain areas dedicated to language, thus assisting patients in restoring access to language and therefore resulting in improvement of language production. Future research should therefore examine whether treatment using inner speech can enhance access to language. Therapy using inner speech has also some major advantages over current therapy techniques. First, for some patients (those who can more easily produce inner than overt speech), it is a less demanding therapy compared to other forms of therapy widely used today. Furthermore, it hinges on the initial level of ability that the patient has, therefore not asking patients to carry out a task that they cannot perform - a request that can cause discouragement and frustration among patients. Secondly, a patient can potentially engage in active therapy without a therapist, by using pre-prepared material. Large amounts of material for therapy can easily be created and personally tailored to the patient's needs and interests. These aspects of therapy using inner speech will inevitably make aphasia therapy more cost effective.

6. Conclusions

This research project examined the behavioural manifestations and neural correlates of inner speech production. Using various methodologies, it was found that young and old healthy participants can easily produce inner speech and use it to complete various language tasks. Patients with aphasia, on the other hand, are a non-homogeneous group in this respect. While some patients can easily produce inner speech, others have great difficulty with the task. Furthermore, the patient data provide evidence of a striking double dissociation between inner and overt speech production.

The imaging results highlight the various components of the language network which are involved in inner speech processing. This includes areas which are highly replicated in language studies (for example, the left IFG and insula), areas which are less commonly implicated with relation to speech production (for example, the angular gyrus), and lastly, non-language regions which support language processing (such as the premotor cortex and the DLPFC).

Like any study, this study had various limitations which are discussed throughout this thesis. Despite these, it is suggested that the results presented here have potential implications for clinical practice, the designing of imaging studies and the conductance of cognitive studies of language processing.

Bibliography

Acheson, D. J. and M. C. MacDonald (2009). "Verbal working memory and language production: Common approaches to the serial ordering of verbal information." Psychological Bulletin **135**(1): 50-68.

Ackermann, H., D. Wildgruber, et al. (1998). "Does the cerebellum contribute to cognitive aspects of speech production? A functional magnetic resonance imaging (fMRI) study in humans." Neuroscience Letters **15**(247): 187-190.

Agosta, F., R. G. Henry, et al. (2010). "Language networks in semantic dementia." Brain **133**: 286-299.

Anderson, J. M., R. Gilmore, et al. (1999). "Conduction aphasia and the arcuate fasciculus: A reexamination of the Wernicke-Geschwind model." Brain and Language **70**(1): 1-12.

Ardila, A. (2010). "A proposed reinterpretation and reclassification of aphasic syndromes." Aphasiology **24**(3): 363-394.

Ardila, A., F. Ostrosky-Solis, et al. (2000). "Age-related cognitive decline during normal aging: The complex effect of education." Archives of Clinical Neuropsychology **15**(6): 495-513.

Ashburner, J. and K. J. Friston (2000). "Voxel-based morphometry - The methods." Neuroimage **11**(6): 805-821.

Ashburner, J. and K. J. Friston (2005). "Unified segmentation." Neuroimage **26**(3): 839-851.

Baayen, R. H., R. Piepenbrock, et al. (1993). The CELEX lexical database (CD-ROM). Philadelphia, Philadelphia: Linguistic Data Consortium.

Bacon, F. (1620). The New Organon: True directions concerning the interpretation of nature. The Works Boston Taggard and Thompson **8**.

Baddeley, A. (1966). "Influence of acoustic and semantic similarity on long-term memory for word sequences " Quarterly Journal of Experimental Psychology **18**: 302.

Baddeley, A. (2003). "Working memory: Looking back and looking forward." Nature Reviews Neuroscience **4**(10): 829-839.

Baddeley, A. and G. Hitch (1974). Working memory. The psychology of learning and motivation: Advances in research and theory. G. H. Bower. New York, Academic Press. **8:** 47-89.

Baddeley, A., N. Thomson, et al. (1975). "Word length and structure of short-term-memory." Journal of Verbal Learning and Verbal Behavior **14**(6): 575-589.

Baddeley, A. and B. Wilson (1985). "Phonological coding and short-term memory in patients without speech." Journal of Memory and Language **24**: 490-502.

Baillieux, H., H. J. De Smet, et al. (2008). "Cerebellar neurocognition: Insights into the bottom of the brain." Clinical Neurology and Neurosurgery **110**(8): 763-773.

Baldo, J. V. and N. F. Dronkers (2007). "Neural correlates of arithmetic and language comprehension: A common substrate?" Neuropsychologia **45**(2): 229-235.

Baldo, J. V., E. C. Klostermann, et al. (2008). "It's either a cook or a baker: Patients with conduction aphasia get the gist but lose the trace." Brain and Language **105**(2): 134-140.

Baldo, J. V., S. Schwartz, et al. (2006). "Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping." Journal of the International Neuropsychological Society **12**(6): 896-900.

Ballet, G. (1886). Le langage interieur et les diverses formes de l'aphasie. Paris, Ancienne Libraire Germer Bailliere.

Barch, D. M., F. W. Sabb, et al. (1999). "Overt verbal responding during fMRI scanning: Empirical investigations of problems and potential solutions." Neuroimage **10**(6): 642-657.

Barton, M. I. (1971). "Recall of generic properties of words in aphasic patients." Cortex **7**(1): 73-82.

Basho, S., E. D. Palmer, et al. (2007). "Effects of generation mode in fMRI adaptations of semantic fluency: Paced production and overt speech." Neuropsychologia **45**(8): 1697-1706.
Basso, A. (2003). Aphasia and its therapy. Oxford, Oxford university press.
Basso, A., M. Gardelli, et al. (1989). "The role of the right-hemisphere in recovery from aphasia - 2 case-studies." Cortex **25**(4): 555-566.
Bates, E., S. M. Wilson, et al. (2003). "Voxel-based lesion-symptom mapping." Nature Neuroscience **6**(5): 448-450.
Belin, P., P. Van Eeckhout, et al. (1996). "Recovery from nonfluent aphasia after melodic intonation therapy: a PET study." Neurology **47**(6): 1504-1511.
Benson, D. F. and A. Ardila (1996). Aphasia: a clinical perspective. Oxford, Oxford University Press.
Berg, E. A. (1948). "A simple objective technique for measuring flexibility in thinking." The Journal of General Psychology **39**: 15-22.
Bhogal, S. K., R. Teasell, et al. (2003). "Intensity of aphasia therapy, impact on recovery." Stroke **34**(4): 987-992.
Biddle, A., L. Watson, et al. (2002). Criteria for determining disability in speech-language disorders. Evidence report/Technology assessment No. 52. Rockville, MD, North Carolina Evidence-Based Practice Center, Agency for Healthcare Research & Quality.
Binder, J. R., R. H. Desai, et al. (2009). "Where Is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies." Cerebral Cortex **19**(12): 2767-2796.
Birn, R. M., P. A. Bandettini, et al. (1999). "Event-related fMRI of tasks involving brief motion." Human Brain Mapping **7**(2): 106-114.
Blonskii, P. P. (1964). Memory and thought. Selected Works in Psychology. Moscow, Prosveshchenie Press
Boatman, D. (2004). "Cortical bases of speech perception: evidence from functional lesion studies." Cognition **92**(1-2): 47-65.
Bookheimer, S. Y., T. A. Zeffiro, et al. (1995). "Regional cerebral blood flow during object naming and word reading." Human Brain Mapping **3**(93-106).
Booth, J. R., D. D. Burman, et al. (2002). "Functional anatomy of intra- and cross-modal lexical tasks." Neuroimage **16**(1): 7-22.
Booth, J. R., S. J. Cho, et al. (2007). "Neural correlates of mapping from phonology to orthography in children performing an auditory spelling task." Developmental Science **10**(4): 441-451.
Borovsky, A., A. P. Saygin, et al. (2007). "Lesion correlates of conversational speech production deficits." Neuropsychologia **45**(11): 2525-2533.
Brassen, S., C. Buchel, et al. (2009). "Structure-function interactions of correct retrieval in healthy elderly women." Neurobiology of Aging **30**(7): 1147-1156.
Brett, M., A. P. Leff, et al. (2001). "Spatial normalization of brain images with focal lesions using cost function masking." Neuroimage **14**(2): 486-500.
Briellmann, R. S., A. Syngeniotis, et al. (2001). "Comparison of hippocampal volumetry at 1.5 tesla and at 3 tesla." Epilepsia **42**(8): 1021-1024.
Broca, P. (1865). "Remarques sur le siège de la faculté du language articulé, suivies d'une observation d'aphémie." Bulletins de la Société Anatomique de Paris **6**: 330-357.
Browman, C. P. and L. Goldstein (1992). "Articulatory phonology - An overview." Phonetica **49**(3-4): 155-180.
Brown, G. D. A. and R. P. W. Loosemore (1994). Computational approaches to normal and impaired spelling. Handbook of Spelling: Theory, Process and Intervention. G. D. A. Brown and N. C. Ellis. Chichester, UK, John Wiley**:** 319-336.
Brown, J. W. (2009). "Inner speech: microgenetic concepts." Aphasiology **23**(5): 531-543.
Brown, R. and D. McNeill (1966). " The "tip of the tongue" phenomenon." Journal of Verbal Learning and Behavior **5**: 325-337.

Bruce, C. and A. Edmundson (2010). "Letting the CAT out of the bag: A review of the Comprehensive Aphasia Test. Commentary on Howard, Swinburn, and Porter, ‘‘Putting the CAT out: What the Comprehensive Aphasia Test has to offer’’." Aphasiology **24**(1): 79-93.
Buchsbaum, B. R. and M. D'Esposito (2008). "The search for the phonological store: From loop to convolution." Journal of Cognitive Neuroscience **20**(5): 762-778.
Buchsbaum, B. R. and M. D'Esposito (2009). "Repetition suppression and reactivation in auditory-verbal short-term recognition memory." Cerebral Cortex **19**(6): 1474-1485.
Bullinaria, J. A. (1994). Connectionist modeling of spelling. Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society. A. Ram and K. Eiselt. Mahwah, Lawrence Erlbaum Assoc Publ**:** 78-83.
Burgess, P. and T. Shallice (1997). The Hayling and Brixton Tests. Bury St Edmunds, UK, Thames Valley Test Company.
Burgess, P. and T. Shallice (1997). The Hayling and Brixton Tests. Test manual. Bury St Edmunds, UK, Thames Valley Test Company.
Burke, D. M., D. G. Mackay, et al. (1991). "On the tip of the tongue - What causes word finding failures in young and older adults." Journal of Memory and Language **30**(5): 542-579.
Burnstine, T. H., R. P. Lesser, et al. (1990). "Characterization of the basal temporal language area in patients with left temporal-lobe epilepsy." Neurology **40**(6): 966-970.
Byng, S., J. Kay, et al. (1990). "Aphasia tests reconsidered." Aphasiology **4**(1): 67-91.
Cabeza, R. (2002). "Hemispheric asymmetry reduction in older adults: The HAROLD model." Psychology and Aging **17**(1): 85-100.
Calvert, G. A., M. J. Brammer, et al. (2000). "Using fMRI to study recovery from acquired dysphasia." Brain and Language **71**(3): 391-399.
Cappa, S. F., D. Perani, et al. (1997). "A PET follow up study of recovery after stroke in acute aphasics." Brain and Language **56**: 55-67.
Cappa, S. F., M. Sandrini, et al. (2002). "The role of the left frontal lobe in action naming - rTMS evidence." Neurology **59**(5): 720-723.
Caramazza, A., A. G. Basili, et al. (1981). "An investigation of repetition and language processing in a case of conduction aphasia." Brain and Language **14**(2): 235-271.
Carreiras, M., M. L. Seghier, et al. (2009). "An anatomical signature for literacy." Nature **461**(7266): 983-U245.
Caspari, I., S. R. Parkinson, et al. (1998). "Working memory and aphasia." Brain and Cognition **37**(2): 205-223.
Catani, M., D. K. Jones, et al. (2005). "Perisylvian language networks of the human brain." Annals of Neurology **57**(1): 8-16.
Catani, M. and M. Mesulam (2008). "The arcuate fasciculus and the disconnection theme in language and aphasia: History and current state." Cortex **44**(8): 953-961.
Cherney, L. R. and S. L. Small (2006). "Task-dependent changes in brain activation following therapy for nonfluent aphasia: Discussion of two individual cases." Journal of the International Neuropsychological Society **12**(6): 828-842.
Christoff, K., V. Prabhakaran, et al. (2001). "Rostrolateral prefrontal cortex involvement in relational integration during reasoning." Neuroimage **14**(5): 1136-1149.
Chumbley, J. R. and K. J. Friston (2009). "False discovery rate revisited: FDR and topological inference using Gaussian random fields." Neuroimage **44**(1): 62-70.
Cohen, L., S. Dehaene, et al. (2000). "The visual word form area - Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients." Brain **123**: 291-307.
Coltheart, M., B. Curtis, et al. (1993). "Models of reading aloud - Dual-route and parallel-distributed-processing approaches." Psychological Review **100**(4): 589-608.
Compston, A. (2006). "On aphasia. By L. Lichtheim, MD, Professor of Medicine in the University of Berne. Brain 1885; 7: 433–484." Brain **129**: 1347-1350.
Conrad, R. (1964). "Acoustic confusions in immediate memory " British Journal of Psychology **55**(1): 75-84.

Corcoran, D. W. J. (1966). "An acoustic factor in letter cancellation." Nature **210**: 658.
Cornelissen, K., M. Laine, et al. (2003). "Adult brain plasticity elicited by anomia treatment." Journal of Cognitive Neuroscience **15**(3): 444-461.
Cotelli, M., R. Manenti, et al. (2006). "Effect of transcranial magnetic stimulation on action naming in patients with Alzheimer disease." Archives of Neurology **63**(11): 1602-1604.
Craik, F. I. M. and E. Bialystok (2006). "Cognition through the lifespan: mechanisms of change." Trends in Cognitive Sciences **10**(3): 131-138.
Cramer, S. C. (2008). "Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery." Annals of Neurology **63**(3): 272-287.
Cramer, S. C., G. Nelles, et al. (1997). "A functional MRI study of subjects recovered from hemiparetic stroke." Stroke **28**(12): 2518-2527.
Cramer, S. C., T. B. Parrish, et al. (2007). "Predicting functional gains in a stroke trial." Stroke **38**(7): 2108-2114.
Crinion, J., J. Ashbumer, et al. (2007). "Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses." Neuroimage **37**(3): 866-875.
Crinion, J. T., E. A. Warburton, et al. (2006). "Listening to narrative speech after aphasic stroke: The role of the left anterior temporal lobe." Cerebral Cortex **16**(8): 1116-1125.
Croquelois, A., M. Wintermark, et al. (2003). "Aphasia in hyperacute stroke: Language follows brain penumbra dynamics." Annals of Neurology **54**: 321-329.
Crosson, B., A. B. Moore, et al. (2005). "Role of the right and left hemispheres in recovery of function during treatment of intention in aphasia." Journal of Cognitive Neuroscience **17**(3): 392-406.
Crosson, B., J. R. Sadek, et al. (2001). "Relative shift in activity from medial to lateral frontal cortex during internally versus externally guided word generation." Journal of Cognitive Neuroscience **13**(2): 272-283.
Cuenod, C. A., S. Y. Bookheimer, et al. (1995). "Functional MRI during word generation, using conventional equipment: A potential tool for language localization in the clinical environment." Neurology **45**(10): 1821-1827.
Curtis, H. S. (1899). "Automatic movement of the larynx." American Journal of Psychology **11**(1): 237-239.
Cusack, R., N. Cumming, et al. (2005). "Automated post-hoc noise cancellation tool for audio recordings acquired in an MRI scanner." Human Brain Mapping **24**(4): 299-304.
Cutler, A. and D. Norris (1988). "The role of strong syllables in segmentation for lexical access." Journal of Experimental Psychology-Human Perception and Performance **14**(1): 113-121.
Dabul, B. L. (1979). Apraxia Battery for Adults (ABA). USA, C.C. Publications, Inc.
De Bleser, R. and J. C. Marshall (2005). "Egon Weigl and the concept of inner speech." Cortex **41**(2): 249-257.
DeLeon, J., R. F. Gottesman, et al. (2007). "Neural regions essential for distinct cognitive porcesses underlying picture naming." Brain **130**: 1408-1422.
Dell, G. S. (1986). "A spreading-activation theory of retrieval in sentence production." Psychological Review **93**(3): 283-321.
Dell, G. S. and P. G. Oseaghdha (1992). "Stages of lexical access in language production." Cognition **42**(1-3): 287-314.
Dell, G. S., M. F. Schwartz, et al. (1997). "Lexical access in aphasic and nonaphasic speakers." Psychological Review **104**(4): 801-838.
Demonet, J. F., G. Thierry, et al. (2005). "Renewal of the neurophysiology of language: Functional neuroimaging." Physiological Reviews **85**(1): 49-95.
Derenzi, E. and L. A. Vignolo (1962). "Token Test - a sensitive test to detect receptive disturbances in aphasics." Brain **85**(4): 665-678.
Devlin, J. T., P. M. Matthews, et al. (2003). "Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study." Journal of Cognitive Neuroscience **15**(1): 71-84.

Dick, F., A. P. Saygin, et al. (2007). "What is involved and what is necessary for complex linguistic and nonlinguistic auditory processing: Evidence from functional magnetic resonance imaging and lesion data." Journal of Cognitive Neuroscience **19**(5): 799-816.

Doesborgh, S. J., M. W. van de Sandt-Koenderman, et al. (2004). "Effects of semantic treatment on verbal communication and linguistic processing in aphasia after stroke: a randomized controlled trial." Stroke **35**(1): 141-146.

Dotson, V. M., M. H. Kitner-Triolo, et al. (2009). "Effects of race and socioeconomic status on the relative influence of education and literacy on cognitive functioning." Journal of the International Neuropsychological Society **15**(4): 580-589.

Dronkers, N. F. (1996). "A new brain region for coordinating speech articulation." Nature **384**(6605): 159-161.

Dronkers, N. F., O. Plaisant, et al. (2007). "Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong." Brain **130**: 1432-1441.

Dronkers, N. F., D. P. Wilkins, et al. (2004). "Lesion analysis of the brain areas involved in language comprehension." Cognition **92**(1-2): 145-177.

Egger, V. (1881). La Parole Interieure. Paris, Ancienne Libraire Germer Bailliere.

Egner, T. and J. Hirsch (2005). "Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information." Nature Neuroscience **8**(12): 1784-1790.

Ellis, A. W. and A. W. Young (1988). Human Cognitive Neuropsychology. London, Lawrence Erlbaum Associates.

Enderby, P. and J. Emerson (1996). "Speech and language therapy: Does it work?" British Medical Journal **312**(7047): 1655-1658.

Etard, O., E. Mellet, et al. (2000). "Picture naming without Broca's and Wernicke's area." Neuroreport **11**(3): 617-622.

Eulitz, C., O. Hauk, et al. (2000). "Electroencephalographic activity over temporal brain areas during phonological encoding in picture naming." Clinical Neurophysiology **111**: 2088-2097.

Feinberg, T. E., L. J. Gonzalez Rothi, et al. (1986). "'Inner speech' in conduction aphasia." Archives of Neurology **43**(6): 591-593.

Fernandez, B., D. Cardebat, et al. (2004). "Functional MRI follow-up study of language processes in healthy subjects and during recovery in a case of aphasia." Stroke **35**(9): 2171-2176.

Fertonani, A., S. Rosini, et al. (2009). "Naming facilitation induced by transcranial direct current stimulation." Behavioral Brain Research **208**(2): 311-318.

Foki, T., A. Gartus, et al. (2008). "Probing overtly spoken language at sentential level - A comprehensive high-field BOLD-fMRI protocol reflecting everyday language demands." Neuroimage **39**(4): 1613-1624.

Folstein, M. F., S. E. Folstein, et al. (1975). "Mini-Mental State - Practical method for grading cognitive state of patients for clinician." Journal of Psychiatric Research **12**(3): 189-198.

Foundas, A. L., S. K. Daniels, et al. (1998). "Anemia: Case studies with lesion localization." Neurocase **4**(1): 35-43.

Frackowiak, R. S. J., C. Weiller, et al. (1991). The functional-anatomy of recovery from brain injury. Exploring Brain Functional Anatomy with Positron Tomography. D. J. Chadwick and J. Whelan. Chichester, John Wiley & Sons Ltd. **163:** 235-244.

Franklin, S., F. Buerk, et al. (2002). "Generalised improvement in speech production for a subject with reproduction conduction aphasia." Aphasiology **16**(10-11): 1087-1114.

Frey, S., J. S. W. Campbell, et al. (2008). "Dissociating the human language pathways with high angular resolution diffusion fiber tractography." Journal of Neuroscience **28**(45): 11435-11444.

Fridriksson, J., L. Morrow-Odom, et al. (2006). "Neural recruitment associated with anomia treatment in aphasia." Neuroimage **32**(3): 1403-1412.

Fridriksson, J., D. Moser, et al. (2007). "Neural correlates of phonological and semantic-based anomia treatment in aphasia." Neuropsychologia **45**(8): 1812-1822.
Friederici, A. D. (2002). "Towards a neural basis of auditory sentence processing." Trends in Cognitive Sciences **6**(2): 78-84.
Friederici, A. D. (2009). "Pathways to language: fiber tracts in the human brain." Trends in Cognitive Sciences **13**(4): 175-181.
Friederici, A. D., J. Bahlmann, et al. (2006). "The brain differentiates human and non-human grammars: Functional localization and structural connectivity." Proceedings of the National Academy of Sciences of the United States of America **103**(7): 2458-2463.
Frings, M., A. Dimitrova, et al. (2006). "Cerebellar involvement in verb generation: An fMRI study." Neuroscience Letters **409**(1): 19-23.
Frost, S. B., S. Barbay, et al. (2003). "Reorganization of remote cortical regions after ischemic brain injury: A potential substrate for stroke recovery." Journal of Neurophysiology **89**(6): 3205-3214.
Furlan, M., G. Marchal, et al. (1996). "Spontaneous neurological recovery after stroke and the fate of the ischemic penumbra." Annals of Neurology **40**: 216-226.
Galdo-Alvarez, S., M. Lindin, et al. (2009). "The effect of age on event-related potentials (ERP) associated with face naming and with the tip-of-the-tongue (TOT) state." Biological Psychology **81**(1): 14-23.
Gauthier, L. V., E. Taub, et al. (2008). "Remodeling the brain - Plastic structural brain changes produced by different motor therapies after stroke." Stroke **39**(5): 1520-1525.
Golestani, N., N. Molko, et al. (2007). "Brain structure predicts the learning of foreign speech sounds." Cerebral Cortex **17**(3): 575-582.
Golestani, N. and C. Pallier (2007). "Anatomical correlates of foreign speech sound production." Cerebral Cortex **17**(4): 929-934.
Golestani, N., T. Paus, et al. (2002). "Anatomical correlates of learning novel speech sounds." Neuron **35**(5): 997-1010.
Goodglass, H. and E. Kaplan (1972). The Assessment of Aphasia and Related Disorders. Philadelphia, Lea & Febiger.
Goodglass, H., E. Kaplan, et al. (1976). "Tip-of-tongue phenomenon in aphasia." Cortex **12**(2): 145-153.
Grady, C. L. (2008). Cognitive neuroscience of aging. Year in Cognitive Neuroscience 2008. Oxford, Blackwell Publishing. **1124:** 127-144.
Grandjean, D., D. Sander, et al. (2008). "Effects of emotional prosody on auditory extinction for voices in patients with spatial neglect." Neuropsychologia **46**(2): 487-496.
Grau-Olivares, M., D. Bartres-Faz, et al. (2007). "Mild cognitive impairment after lacunar infarction: Voxel-based morphometry and neuropsychological assessment." Cerebrovascular Diseases **23**(5-6): 353-361.
Greener, J., P. Enderby, et al. (2000). "Speech and language therapy for aphasia following stroke." Cochrane Database of Systematic Reviews **2**: CD000425.
Grogan, A., D. W. Green, et al. (2009). "Structural correlates of semantic and phonemic fluency ability in first and second languages." Cerebral Cortex **19**(11): 2690-2698.
Guenther, F. H. (2005). Cortical interactions underlying the production of speech sounds. Annual Meeting of the American-Speech-Language-Hearing-Association, San Diego, CA, Elsevier Science Inc.
Harding, D. and C. Pound (1999). Needs, function and measurement: Juggling with multiple language impairments. The Aphasia Therapy File. S. Byng, K. Swinburn and C. Pound. East Sussex, UK, Psychology Press. **1:** 13-40.
Hardyck, C. D. and L. Petrinov (1970). "Subvocal speech and comprehension level as a ffunction of difficulty level of reading material." Journal of Verbal Learning and Verbal Behavior **9**(6): 647-652.
Harris, K. C., J. R. Dubno, et al. (2009). "Speech recognition in younger and older adults: A dependency on low-level auditory cortex." Journal of Neuroscience **29**(19): 6078-6087.

Head, H. (1926). Aphasia and Kindred Disorders of Speech. London, Cambridge University Press.
Hebert, S., A. Racette, et al. (2003). "Revisiting the dissociation between singing and speaking in expressive aphasia." Brain **126**(Pt 8): 1838-1850.
Heine, M. K., B. A. Ober, et al. (1999). "Naturally occurring and experimentally induced tip-of-the-tongue experiences in three adult age groups." Psychology and Aging **14**(3): 445-457.
Heiss, W. D. and A. Thiel (2006). "A proposed regional hierarchy in recovery of post-stroke aphasia." Brain and Language **98**(1): 118-123.
Heiss, W. D., A. Thiel, et al. (2003). "Disturbance and recovery of language function: Correlates in pet activation studies." Neuroimage **20**(Suppl 1): S42-S49.
Henry, M. L., P. M. Beeson, et al. (2007). "The role of left perisylvian cortical regions in spelling." Brain and Language **100**(1): 44-52.
Hickok, G., P. Erhard, et al. (2000). "a functional magnetic resonance imaging study of the role of the left posterior superior temporal gyrus in speech production: implications for the explanation of conduction aphasia." Neuroscience Letters **287**: 156-160.
Hickok, G. and D. Poeppel (2007). "Opinion - The cortical organization of speech processing." Nature Reviews Neuroscience **8**(5): 393-402.
Hillis, A. E., P. B. Barker, et al. (2000). "MR perfusion imaging reveals regions of hypoperfusion associated with aphasia and neglect." Neurology **55**(6): 782-788.
Hillis, A. E., J. T. Kleinman, et al. (2006). "Restoring cerebral blood flow reveals neural regions critical for naming " The Journal of Neuroscience **26**(31): 8069-8073.
Hillis, A. E., R. J. Wityk, et al. (2002). "Subcortical aphasia and neglect in acute stroke: The role of cortical hypoperfusion." Brain **125**: 1094-1104.
Hillis, A. E., M. Work, et al. (2004). "Re-examining the brain regions crucial for orchestrating speech articulation." Brain **127**(Pt 7): 1479-1487.
Hoeft, F., A. Meyler, et al. (2007). "Functional and morphometric brain dissociation between dyslexia and reading ability." Proceedings of the National Academy of Sciences of the United States of America **104**(10): 4234-4239.
Holland, A. L. (1980). Communicative Abilities in Daily Living. Baltimore, University Park Press.
Houghton, G. and M. Zorzi (2003). "Normal and impaired spelling in a connectionist dual-route architecture." Cognitive Neuropsychology **20**(2): 115-162.
Howard, D. and S. Franklin (1990). Memory without rehearsal. Neuropsychological impairments of short-term memory. G. Vallar and T. Shallice. Cambridge, UK, Cambridge University Press**:** 287-318.
Howard, D. and F. M. Hatfield (1987). Aphasia Therapy; Historical and Contemporary Issues. London, Lawrence Erlbaum.
Howard, D., K. Patterson, et al. (1984). Variability and consistency in picture naming by aphasic patients. Recent advances in neurology. F. C. Rose. New York, Raven Press. **42: progress in aphasiology:** 263-276.
Howard, D., K. Swinburn, et al. (2010). "Putting the CAT out: What the Comprehensive Aphasia Test has to offer." Aphasiology **24**(1): 56-74.
Huang, J., T. H. Carr, et al. (2002). "Comparing cortical activations for silent and overt speech using event-related fMRI." Human Brain Mapping **15**(1): 39-53.
Huber, W., K. Poeck, et al. (1984). The Aachen Aphasia Test. Advances in Neurology 42; Progress in Aphasiology. F. C. Rose. New York, Raven Press. **42:** 291-303.
Huettig, F. and R. J. Hartsuiker (2009). "Listening to yourself is like listening to others: External, but not internal, verbal self-monitoring is based on speech perception." Language and Cognitive Processes.
Hughlings Jackson, J. (1879). "On affectations of speech from disease of the brain." Brain **1**: 304-330.
Hulme, C., S. Roodenrys, et al. (1997). "Word-frequency effects on short-term memory tasks: Evidence for a redintegration process in immediate serial recall." Journal of Experimental Psychology-Learning Memory and Cognition **23**(5): 1217-1232.

Hunter, S. M., P. Crome, et al. (2006). "Development of treatment schedules for research: a structured review to identify methodologies used and a worked example of 'mobilisation and tactile stimulation' for stroke patients." Physiotherapy **92**(4): 195-207.
Iglesias, S., G. Marchal, et al. (1996). "Do changes in oxygen metabolism in the unaffected cerebral hemisphere underlie early neurological recovery after stroke? A positron emission tomography study." Stroke **27**: 1192-1199.
Indefrey, P. and W. J. M. Levelt (2004). "The spatial and temporal signatures of word production components." Cognition **92**(1-2): 101-144.
Ivnik, R. J., J. F. Malec, et al. (1996). "Neuropsychological tests' norms above age 55: COWAT, BNT, MAE token, WRAT-R reading, AMNART, STROOP, TMT, and JLO." Clinical Neuropsychologist **10**(3): 262-278.
James, W. (1890). The Principles of Psychology. New York, H. Holt.
Jancke, L. and N. J. Shah (2004). "Hearing syllables by seeing visual stimuli." European Journal of Neuroscience **19**(9): 2603-2608.
Jeannerod, M. (1995). "Mental imagery in the motor context." Neuropsychologia **33**(11): 1419-1432.
Justus, T. (2004). "The cerebellum and English grammatical morphology: Evidence from production, comprehension, and grammaticality judgments." Journal of Cognitive Neuroscience **16**(7): 1115-1130.
Kaplan, E. F., H. Goodglass, et al. (2001). Boston Naming Test (BNT), Lippincott Williams & Wilkins.
Katanoda, K., K. Yoshikawa, et al. (2001). "A functional MRI study on the neural substrates for writing." Human Brain Mapping **13**(1): 34-42.
Kay, J., M. Coltheart, et al. (1992). Psycholinguistic Assessment of Language Processing, Psychology Press.
Keller, S. S., T. Crow, et al. (2009). "Broca's area: Nomenclature, anatomy, typology and asymmetry." Brain and Language **109**(1): 29-48.
Kelly, M. H. and J. K. Bock (1988). "Stress in time." Journal of Experimental Psychology-Human Perception and Performance **14**(3): 389-403.
Kelso, J. A. S. (1982). Concepts and issues in human motor behavior: Coming to grips wtih the jargon. Human Motor Behavior: An Introdcution. J. A. S. Kelso, Lawrence Erlbaum Associates Inc**:** 21-62.
Kerns, J. G., J. D. Cohen, et al. (2004). "Anterior Cingulate conflict monitoring and adjustments in control." Science **303**(5660): 1023-1026.
Kertesz, A. (1982). The Western Aphasia Battery. New York, Grune & Stratton.
Kim, J. N., Y. W. Kim, et al. (2006). "Diffusion-tensor magnetic resonance imaging in children with language impairment." Neuroreport **17**(12): 1279-1282.
Kinsbourne, M. (2000). "Inner speech and the inner life." Brain and Language **71**(1): 120-123.
Koppel, S. and C. Buchel (2005). "Alternatives to the Wada test: a critical view of functional magnetic resonance imaging in preoperative use." Current Opinion in Neurology **18**(4): 418-423.
Kriegeskorte, N., W. K. Simmons, et al. (2009). "Circular analysis in systems neuroscience: the dangers of double dipping." Nature Neuroscience **12**(5): 535-540.
Kuljic-Obradovic, D. C. (2003). "Subcortical aphasia: Three different language disorder syndromes?" European Journal of Neurology **10**(4): 445-448.
Kurland, J., M. A. Naeser, et al. (2004). "Test-retest reliability of fMRI during nonverbal semantic decisions in moderate-severe nonfluent aphasia patients." Behavioural Neurology **15**(3-4): 87-97.
Lambon-Ralph, M. A., G. Pobric, et al. (2009). "Conceptual knowledge is underpinned by the temporal pole bilaterally: convergent evidence from rTMS." Cerebral Cortex **19**(4): 832-838.
Lashley, K. S. (1929). Brain Mechanisms and Intelligence. Chicago, University of Chicago Press.

Lazar, R. M. and D. Antoniello (2008). "Variability in recovery from aphasia." Current Neurology and Neuroscience Reports **8**(6): 497-502.
Ledorze, G., N. Boulay, et al. (1994). "The contrasting effects of a semantic versus a formal semantic technique for the facilitation of naming in a case of anomia." Aphasiology **8**(2): 127-141.
Lee, H., T. Nakada, et al. (1984). "Transfer of language dominance." Annals of Neurology **15**(3): 304-307.
Leff, A. P., T. M. Schofield, et al. (2009). "The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: evidence from 210 patients with stroke." Brain **132**(12): 3401-3410.
Leger, A., J. F. Demonet, et al. (2002). "Neural substrates of spoken language rehabilitation in an aphasic patient: An fMRI study." Neuroimage **17**(1): 174-183.
Levelt, W. J. M. (1983). "Monitoring and self-repair in speech." Cognition **14**(1): 41-104.
Levelt, W. J. M. (1989). Speaking: From Intention to Articulation. Cambridge, MA, MIT Press.
Levelt, W. J. M. (1993). Speaking: From Intention to Articulation. Cambridge, MA, MIT Press.
Levelt, W. J. M. (1999). "Models of word production." Trends in Cognitive Sciences **3**(6): 223-232.
Levelt, W. J. M., P. Praamstra, et al. (1998). "An MEG study of picture naming." Journal of Cognitive Neuroscience **10**(5): 553-567.
Levelt, W. J. M., A. Roelofs, et al. (1999). "A theory of lexical access in speech production." The Behavioral and Brain Sciences **22**(1): 1-75.
Levelt, W. J. M. and L. Wheeldon (1994). "Do speakers have access to a mental syllabary." Cognition **50**(1-3): 239-269.
Levine, D. N., R. Calvanio, et al. (1982). "Language in the absence of inner speech." Neuropsychologia **20**(4): 391-409.
Levine, D. N. and J. P. Mohr (1979). "Language after bilateral cerebral infarctions: Role of the minor hemisphere in speech." Neurology **29**(7): 927-938.
Lichtheim, L. (1885). "On aphasia." Brain **7**: 433-484.
Liu, C. L., C. W. Hue, et al. (2006). "Dissociated roles of the middle frontal gyri in the processing of Chinese characters." Neuroreport **17**(13): 1397-1401.
Luaute, J., P. Halligan, et al. (2006). "Visuo-spatial neglect: A systematic review of current interventions and their effectiveness." Neuroscience and Biobehavioral Reviews **30**(7): 961-982.
Luders, H., R. P. Lesser, et al. (1986). "Basal temporal language area demonstrated by electrical-stimulation." Neurology **36**(4): 505-510.
Luders, H., R. P. Lesser, et al. (1991). "Basal temporal language area." Brain **114**: 743-754.
Lum, C. and A. W. Ellis (1994). "Is ‘nonpropositional’ speech preserved in aphasia?" Brain and Language **46**: 368–391.
Lum, C. and A. W. Ellis (1999). "Why do some aphasics show an advantage on some tests of nonpropositional (automatic) speech?" Brain and Language **70**(1): 95-118.
Luria, A. R. (1963). The Human Brain and Psychic Processes. Moscow, Academy of Pedagogical Sciences, RSFSR.
Luria, A. R. (1964). Factors and Forms of Aphasia. Ciba Foundation Symposium - Disorders of Language. A. V. S. de Reuck and M. O'Connor. London, Ciba Foundation**:** 143-161.
Luria, A. R. (1976). Basic Problems of Neurolinguistics. The Hague, Mouton.
Lurito, J. T., D. A. Kareken, et al. (2000). "Comparison of rhyming and word generation with FMRI." Human Brain Mapping **10**(3): 99-106.
Manly, J. J., D. M. Jacobs, et al. (1999). "Quality of education and neuropsychological test performance among nondemented community-dwelling elders." Neurology **52**(6): A433-A433.

Mansouri, F. A., K. Tanaka, et al. (2009). "Conflict-induced behavioural adjustment: A clue to the executive functions of the prefrontal cortex." Nature Reviews Neuroscience **10**(2): 141-152.
Mao, H., S. H. Polensek, et al. (2007). "Diffusion tensor and functional magnetic resonance imaging of diffuse axonal injury and resulting language impairment." Journal of Neuroimaging **17**(4): 292-294.
Marcopulos, B. A., C. A. McLain, et al. (1997). "Cognitive impairment or inadequate norms? A study of healthy, rural, older adults with limited education." Clinical Neuropsychologist **11**(2): 111-131.
Marien, P., S. Engelborghs, et al. (2000). "Aphasia following cerebellar damage: Fact or fallacy?" Journal of Neurolinguistics **13**(2-3): 145-171.
Marien, P., E. Mampaey, et al. (1998). "Normative data for the Boston Naming Test in native Dutch-speaking Belgian elderly." Brain and Language **65**(3): 447-467.
Marien, P., J. Saerens, et al. (1996). "Cerebellar induced aphasia: Case report of cerebellar induced prefrontal aphasic language phenomena supported by SPECT findings." Journal of the Neurological Sciences **144**(1-2): 34-43.
Marsh, E. B. and A. E. Hillis (2006). Recovery from aphasia following brain injury: The role of reorganization. Reprogramming the Brain. A. R. Moller. Amsterdam, Elsevier Science Bv. **157:** 143-156.
Marshall, J. (2010). "Classification of aphasia: Are there benefits for practice?" Aphasiology **24**(3): 408-412.
Marshall, J., J. Robson, et al. (1998). "Why does monitoring fail in jargon aphasia? Comprehension, judgment, and therapy evidence." Brain and Language **63**(1): 79-107.
Marshall, J. C. and F. Newcombe (1973). "Patterns of paralexia: A psycholinguistic approach." Journal of Psycholinguistic Research **2**: 175–199.
Marshall, R. C., S. I. Neuburger, et al. (1994). "Verbal self-correction and improvement in treated aphasic clients." Aphasiology **8**(6): 535-547.
Marshall, R. C., B. Z. Rappaport, et al. (1985). "Self-monitoring behavior in a case of severe auditory agnosia with aphasia." Brain and Language **24**(2): 297-313.
Marshall, R. S., E. Zarahn, et al. (2009). "Early imaging correlates of subsequent motor recovery after stroke." Annals of Neurology **65**(5): 596-602.
Martin, N., G. S. Dell, et al. (1994). "Origins of paraphasias in deep dysphasia - Testing the consequences of a decay impairment to an interactive spreading activation model of lexical retrieval." Brain and Language **47**(4): 609-660.
Martin, N. and E. M. Saffran (1997). "Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes." Cognitive Neuropsychology **14**(5): 641-682.
Martin, N., E. M. Saffran, et al. (1996). "Recovery in deep dysphasia: Evidence for a relation between auditory-verbal STM capacity and lexical errors in repetition." Brain and Language **52**(1): 83-113.
Martin, R. C. (2003). "Language processing: Functional organization and neuroanatomical basis." Annual Review of Psychology **54**: 55-89.
Martin, R. C. and A. Caramazza (1982). "Short-term memory performance in the absence of phonological coding." Brain and Cognition **1**(50-70).
Martin, R. C., M. F. Lesch, et al. (1999). "Independence of input and output phonology in word processing and short-term memory." Journal of Memory and Language **41**(1): 3-29.
Mashal, N., M. Faust, et al. (2009). "An fMRI study of processing novel metaphoric sentences." Laterality **14**(1): 30-54.
Mathuranath, P. S., P. J. Nestor, et al. (2000). "A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia." Neurology **55**: 1613-1620.
Matsumoto, R., D. R. Nair, et al. (2004). "Functional connectivity in the human language system: a cortico-cortical evoked potential study." Brain **127**: 2316-2330.

McCarthy, G., A. M. Blamire, et al. (1993). "Echo-planar magnetic resonance imaging studies of frontal cortex activation during word generation in humans." Proceedings of the National Academy of Sciences of the United States of America **90**: 4952-4956.

McDermott, K. B., S. E. Petersen, et al. (2003). "A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging." Neuropsychologia **41**(3): 293-303.

McGuire, P. K., D. A. Silbersweig, et al. (1996). "Functional anatomy of inner speech and auditory verbal imagery." Psychological Medicine **26**(1): 29-38.

Mechelli, A., C. J. Price, et al. (2005). "Voxel-based morphometry of the human brain: Methods and applications." Current Medical Imaging Reviews **1**(2): 105-113.

Meinzer, M. and C. Breitenstein (2008). "Functional imaging studies of treatment-induced recovery in chronic aphasia." Aphasiology **22**(12): 1251-1268.

Meinzer, M., T. Elbert, et al. (2004). "Intensive language training enhances brain plasticity in chronic aphasia." BMC Biology **2**: 20.

Meinzer, M., T. Flaisch, et al. (2008). "Functional re-recruitment of dysfunctional brain areas predicts language recovery in chronic aphasia." Neuroimage **39**(4): 2038-2046.

Meinzer, M., T. Flaisch, et al. (2006). "Brain regions essential for improved lexical access in an aged aphasic patient: a case report." BMC Neurology **6**: 10.

Meinzer, M., T. Flaisch, et al. (2009). "Neural signatures of semantic and phonemic fluency in young and old adults." Journal of Cognitive Neuroscience **21**(10): 2007-2018.

Meinzer, M., J. Obleser, et al. (2007). "Recovery from aphasia as a function of language therapy in an early bilingual patient demonstrated by fMRI." Neuropsychologia **45**(6): 1247-1256.

Meyers, J. E. and K. R. Meyers (1995). Rey Complex Figure Test and Recognition Trial, Psychological Assessment Resource, Inc.

Miceli, G., A. Amitrano, et al. (1996). "The treatment of anomia resulting from output lexical damage: Analysis of two cases." Brain and Language **52**: 150-174.

Mitrushina, M. and P. Satz (1995). "Repeated testing of normal elderly with the Boston Naming Test." Aging-Clinical and Experimental Research **7**(2): 123-127.

Miyake, A., P. A. Carpenter, et al. (1994). "A capacity approach to syntactic comprehension disorders - Making normal adults perform like aphasic patients." Cognitive Neuropsychology **11**(6): 671-717.

Molenberghs, P., C. R. Gillebert, et al. (2008). "Convergence between lesion-symptom mapping and functional magnetic resonance imaging of spatially selective attention in the intact brain." Journal of Neuroscience **28**(13): 3359-3373.

Monsell, S. (1987). On the relation between lexical input and output pathways for speech. Language Perception and Production. A. Allport, D. G. Mackay, W. Prinz and E. Scheerer. London Academic Press**:** 273–311.

Moore, C. J. and C. J. Price (1999). "Three distinct ventral occipitotemporal regions for reading and object naming." Neuroimage **10**(2): 181-192.

Mori, S. and P. C. M. van Zijl (2002). "Fiber tracking: Principles and strategies - A technical review." NMR in Biomedicine **15**(7-8): 468-480.

Mueller, N. G. and R. T. Knight (2006). "The functional neuroanatomy of working memory: Contributions of human brain lesion studies." Neuroscience **139**(1): 51-58.

Mummery, C. J., K. Patterson, et al. (1996). "Generating 'tiger' as an animal name or a word beginning with T: Differences in brain activation." Proceedings of the Royal Society of London Series B-Biological Sciences **263**(1373): 989-995.

Naeser, M. A., P. I. Martin, et al. (2004). "Overt propositional speech in chronic nonfluent aphasia studied with the dynamic susceptibility contrast fMRI method." Neuroimage **22**(1): 29-41.

Naeser, M. A., P. I. Martin, et al. (2005). "Improved naming after TMS treatments in a chronic, global aphasia patient - case report." Neurocase **11**(3): 182-193.

Nebes, R. D. (1975). "The nature of internal speech in a patient with aphemia." Brain and Language **2**: 489-497.

Nettleton, J. and R. Lesser (1991). "Therapy for naming difficulties in aphasia: Application of a cognitive neuropsychological model." Journal of Neurolinguistics **6**: 139-157.
Nickels, L. (2002). "Improving word finding: Practice makes (closer to) perfect?" Aphasiology **16**(10-11): 1047-1060.
Nickels, L. (2002). "Therapy for naming disorders: Revisiting, revising, and reviewing." Aphasiology **16**(10-11): 935-979.
Nickels, L. and D. Howard (1995). "Phonological errors in aphasic naming - comprehension, monitoring and lexicality." Cortex **31**(2): 209-237.
Nolan, K. A. and A. Caramazza (1982). "Modality-independent impairments in word-processing in a deep dyslexic patient." Brain and Language **16**(2): 237-264.
Noppeney, U., K. Patterson, et al. (2007). "Temporal lobe lesions and semantic impairment: A comparison of herpes simplex virus encephalitis and semantic dementia." Brain **130**: 1138-1147.
Nudo, R. J., B. M. Wise, et al. (1996). "Neural substrates for the effects of rehabilitative training on motor recovery after ischemic infarct." Science **272**(5269): 1791-1794.
Obleser, J. and S. A. Kotz (2010). "Expectancy constraints in degraded speech modulate the language comprehension network." Cerebral Cortex **20**(3): 633-640.
Ojemann, J. G., R. L. Buckner, et al. (1998). "Functional MRI studies of word-stem completion: Reliability across laboratories and comparison to blood flow imaging with PET." Human Brain Mapping **6**: 203-215.
Okada, K., K. R. Smith, et al. (2003). "Word length modulates neural activity in auditory cortex during covert object naming." Neuroreport **14**(18): 2323-2326.
Oldfield, R. C. (1971). "The assessment and analysis of handedness: The Edinburgh Inventory." Neuropsychologia **1**(9): 97-113.
Olson, A. and A. Caramazza (1994). Representation and connectionist models: The NETspell experience. Handbook of Spelling: Theory, Process and Intervention. G. D. A. Brown and N. C. Ellis. Chichester, UK, John Wiley**:** 337-364.
Omura, K., T. Tsukamoto, et al. (2004). "Neural correlates of phoneme-to-grapheme conversion." Neuroreport **15**(6): 949-953.
Oomen, C. C., A. Postma, et al. (2001). "Prearticulatory and postarticulatory self-monitoring in Broca's aphasia." Cortex **37**(5): 627-641.
Oppenheim, G. M. and G. S. Dell (2008). "Inner speech slips exhibit lexical bias, but not the phonemic similarity effect." Cognition **106**(1): 528-537.
Owen, A. M., K. M. McMillan, et al. (2005). "N-back working memory paradigm: A meta-analysis of normative functional neuroimaging." Human Brain Mapping **25**(1): 46-59.
Owen, W. J., R. Borowsky, et al. (2004). "FMRI of two measures of phonological processing in visual word recognition: Ecological validity matters." Brain and Language **90**(1-3): 40-46.
Ozdemir, R., A. Roelofs, et al. (2007). "Perceptual uniqueness point effects in monitoring internal speech." Cognition **105**(2): 457-465.
Paap, K. R. and R. W. Noel (1991). "Dual-route models of print to sound - Still a good horse race." Psychological Research-Psychologische Forschung **53**(1): 13-24.
Page, M. P. A., A. Madge, et al. (2007). "Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus." Journal of Memory and Language **56**(1): 49-64.
Palmer, E. D., H. J. Rosen, et al. (2001). "An event-related fMRI study of overt and covert word stem completion." Neuroimage **14**: 182-193.
Papathanassiou, D., O. Etard, et al. (2000). "A common language network for comprehension and production: A contribution to the definition of language epicenters with PET." Neuroimage **11**(4): 347-357.
Parker, G. J. M., S. Luzzi, et al. (2005). "Lateralization of ventral and dorsal auditory-language pathways in the human brain." Neuroimage **24**(3): 656-666.
Patterson, K., P. J. Nestor, et al. (2007). "Where do you know what you know? The representation of semantic knowledge in the human brain." Nature Reviews Neuroscience **8**(12): 976-987.

Patterson, K. E. and J. Morton (1985). From orthography to phonology: An attempt at an old interpretation. Surface Dyslexia. K. E. Patterson, J. Marshall and M. Coltheart. London, Lawrence Erlbaum Associates**:** 335-359.
Paulesu, E., C. D. Frith, et al. (1993). "The neural correlates of the verbal component of working memory." Nature **362**(6418): 342-345.
Peck, K. K., A. B. Moore, et al. (2004). "Functional magnetic resonance imaging before and after aphasia therapy - Shifts in hemodynamic time to peak during an overt language task." Stroke **35**(2): 554-559.
Perani, D., S. F. Cappa, et al. (2003). "A fMRI study of word retrieval in aphasia." Brain and Language **85**: 357-368.
Petersen, S. E., P. T. Fox, et al. (1989). "Positron emission tomographic studies of the processing of single words." Journal of Cognitive Neuroscience **1**: 153-170.
Petrides, M. and D. N. Pandya (2009). "Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey." Plos Biology **7**(8): 16.
Piras, F. and P. Marangolo (2007). "Noun-verb naming in aphasia: A voxel-based lesion-symptom mapping study." Neuroreport **18**(14): 1455-1458.
Plaut, D. C., J. L. McClelland, et al. (1996). "Understanding normal and impaired word reading: Computational principles in quasi-regular domains." Psychological Review **103**(1): 56-115.
Ploner, C. J., B. M. Gaymard, et al. (2005). "The prefrontal substrate of reflexive saccade inhibition in humans." Biological Psychiatry **57**(10): 1159-1165.
Poeppel, D. and G. Hickok (2004). "Towards a new functional anatomy of language." Cognition **92**(1-2): 1-12.
Poldrack, R. A., E. Temple, et al. (2001). "Relations between the neural bases of dynamic auditory processing and phonological processing: Evidence from fMRI." Journal of Cognitive Neuroscience **13**(5): 687-697.
Poldrack, R. A., A. D. Wagner, et al. (1999). "Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex." Neuroimage **10**(1): 15-35.
Porch, B. E. (1971). Porch Index of Communicative Ability. Palo Alto, CA, Consulting Psychologists Press.
Postma, A. (2000). "Detection of errors during speech production: A review of speech monitoring models." Cognition **77**(2): 97-131.
Postma, A. and C. Noordanus (1996). "Production and detection of speech errors in silent, mouthed, noise-masked, and normal auditory feedback speech." Language and Speech **39**: 375-392.
Pratap, R. C. (1987). "The speech evoked potential in normal subjects and patients with cerebral hemispheric lesions." Clinical Neurology and Neurosurgery **89**(4): 237-42.
Price, C. J. (2000). "The anatomy of language: Contributions from functional neuroimaging." Journal of Anatomy **197**: 335-359.
Price, C. J. (2010). "The anatomy of language: A review of 100 fMRI studies published in 2009." Annals of the New York Academy of Science **1191**(1): 62-88.
Price, C. J. and J. Crinion (2005). "The latest on functional imaging studies of aphasic stroke." Current Opinion in Neurology **18**(4): 429-434.
Price, C. J., J. Crinion, et al. (2006). "Design and analysis of fMRI studies with neurologically impaired patients." Journal of Magnetic Resonance Imaging **23**(6): 816-826.
Price, C. J. and J. T. Devlin (2004). "The pro and cons of labelling a left occipitotemporal region: "The visual word form area"." Neuroimage **22**(1): 477-479.
Price, C. J., J. T. Devlin, et al. (2005). "Meta-analyses of object naming: Effect of baseline." Human Brain Mapping **25**(1): 70-82.
Price, C. J. and K. J. Friston (1997). "Cognitive conjunction: A new approach to brain activation experiments." Neuroimage **5**(4): 261-270.
Price, C. J. and K. J. Friston (1999). "Scanning patients with tasks they can perform." Human Brain Mapping **8**(2-3): 102-108.

Price, C. J., C. J. Mummery, et al. (1999). "Delineating necessary and sufficient neural systems with functional imaging studies of neuropsychological patients." Journal of Cognitive Neuroscience **11**(4): 371-382.
Price, C. J., M. L. Seghier, et al. (2010). "Predicting language outcome and recovery after stroke: The PLORAS system." Nature Reviews Neurology **6**(4): 202-210.
Price, C. J., E. A. Warburton, et al. (2001). "Dynamic diaschisis: Anatomically remote and context-sensitive human brain lesions." Journal of Cognitive Neuroscience **13**(4): 419-429.
Pugh, K. R., W. E. Mencl, et al. (2000). "Functional neuroimaging studies of reading and reading disability (developmental dyslexia)." Mental Retardation and Developmental Disabilities Research Reviews **6**(3): 207-213.
Pugh, K. R., B. A. Shaywitz, et al. (1996). "Cerebral organization of component processes in reading." Brain **119**: 1221-1238.
Pulvermuller, F., O. Hauk, et al. (2005). "Therapy-related reorganization of language in both hemispheres of patients with chronic aphasia." Neuroimage **28**(2): 481-489.
Quigg, M., D. S. Geldmacher, et al. (2006). "Conduction aphasia as a function of the dominant posterior perisylvian cortex - Report of two cases." Journal of Neurosurgery **104**(5): 845-848.
Raboyeau, G., X. De Boissezon, et al. (2008). "Right hemisphere activation in recovery from aphasia - Lesion effect or function recruitment?" Neurology **70**(4): 290-298.
Rapcsak, S. Z., P. M. Beeson, et al. (2009). "Phonological dyslexia and dysgraphia: Cognitive mechanisms and neural substrates." Cortex **45**(5): 575-591.
Rapp, B. and M. Goldrick (2006). "Speaking words: Contributions of cognitive neuropsychological research." Cognitive Neuropsychology **23**(1): 39-73.
Rastle, K. and M. Coltheart (2000). "Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords." Journal of Memory and Language **42**(3): 342-364.
Raven, J. C. (1938). Raven's Progressive Matrices, Harcourt Assessment.
Raymer, A. M., A. L. Foundas, et al. (1997). "Cognitive neuropsychological analysis and neuroanatomic correlates in a case of acute anomia." Brain and Language **58**(1): 137-156.
Richardson, F. M. and C. J. Price (2009). "Structural MRI studies of language function in the undamaged brain." Brain Structure & Function **213**(6): 511-523.
Richter, M., W. H. R. Miltner, et al. (2008). "Association between therapy outcome and right-hemispheric activation in chronic aphasia." Brain **131**: 1391-1401.
Richter, S., M. Gerwig, et al. (2007). "Cognitive functions in patients with MR-defined chronic focal cerebellar lesions." Journal of Neurology **254**(9): 1193-1203.
Rijntjes, M. (2006). "Mechanisms of recovery in stroke patients with hemiparesis or aphasia: new insights, old questions and the meaning of therapies." Current Opinion in Neurology **19**(1): 76-83.
Robey, R. R. (1998). "A meta-analysis of clinical outcomes in the treatment of aphasia." Journal of Speech Language and Hearing Research **41**(1): 172-187.
Ross, P. (1983). "Phonological processing during silent reading in aphasic patients." Brain and Language **19**(2): 191-203.
Rowan, A., F. Vargha-Khadem, et al. (2007). "Cortical abnormalities and language function in young patients with basal ganglia stroke." Neuroimage **36**(2): 431-440.
Rubin, D. C. (1975). "Within word structure in the tip-of-the-tongue phenomenon." Journal of Verbal Learning and Verbal Behavior **14**(4): 392-397.
Ryding, E., B. Bradvik, et al. (1996). "Silent speech activates prefrontal cortical regions asymmetrically, as well as speech-related areas in the dominant hemisphere." Brain and Language **52**(3): 435-451.
Saffran, E. M. and O. S. M. Marin (1975). "Immediate memory for word lists and sentences in a patient with deficient auditory short-term-memory." Brain and Language **2**(4): 420-433.
Salthouse, T. A. (2003). Interrelations of aging, knowledge, and cognitive performance. Understanding Human Development: Dialogues with Lifespan Psychology. U. M.

Staudinger and U. Lindenberger. Norwell, Massachusetts, Kluwer Academic Publishers**:** 265–287.
Sarno, M. T. (1969). The Functional Communication Profile. New York, Institute of Rehabilitation Medicine, New York University Medical Center.
Sato, M., M. Baciu, et al. (2004). "Multistable representation of speech forms: A functional MRI study of verbal transformations." Neuroimage **23**(3): 1143-1151.
Saur, D., B. W. Kreher, et al. (2008). "Ventral and dorsal pathways for language." Proceedings of the National Academy of Sciences of the United States of America **105**(46): 18035-18040.
Saur, D., R. Lange, et al. (2006). "Dynamics of language reorganization after stroke." Brain **129**: 1371-1384.
Saur, D., O. Ronneberger, et al. (2010). "Early functional magnetic resonance imaging activations predict language outcome after stroke." Brain **133**(1252-1264).
Saur, D., B. Schelter, et al. (2009). "Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension " Neuroimage **49**(4): 3187-3197
Saygin, A. P. (2007). "Superior temporal and premotor brain areas necessary for biological motion perception." Brain **130**: 2452-2461.
Saygin, A. P., S. M. Wilson, et al. (2004). "Action comprehension in aphasia: Linguistic and non-linguistic deficits and their lesion correlates." Neuropsychologia **42**(13): 1788-1804.
Schlaggar, B. L. and B. D. McCandliss (2007). "Development of neural systems for reading." Annual Review of Neuroscience **30**: 475-503.
Schlenck, K. J., W. Huber, et al. (1987). ""Prepairs" and repairs: Different monitoring functions in aphasic language production." Brain and Language **30**(2): 226-44.
Schoch, B., A. Dimitrova, et al. (2006). "Functional localization in the human cerebellum based on voxelwise statistical analysis: A study of 90 patients." Neuroimage **30**(1): 36-51.
Schuell, H. M. (1965). Minnesota Test for Differential Diagnosis of Aphasia. Minneapolis, University of Minnesota Press.
Schwartz, M. F., D. Y. Kimberg, et al. (2009). "Anterior temporal involvement in semantic word retrieval: Voxel-based lesion-symptom mapping evidence from aphasia." Brain **132**(12): 3411-3427.
Seidenberg, M. S. and J. L. McClelland (1989). "A distributed, developmental model of word recognition and naming." Psychological Review **96**(4): 523-568.
Selnes, O. A., P. C. M. van Zijl, et al. (2002). "MR diffusion tensor imaging documented arcuate fasciculus lesion in a patient with normal repetition performance." Aphasiology **16**(9): 897-901.
Shafto, M. A., D. M. Burke, et al. (2007). "On the tip-of-the-tongue: Neural correlates of increased word-finding failures in normal aging." Journal of Cognitive Neuroscience **19**(12): 2060-2070.
Shafto, M. A., E. A. Stamatakis, et al. (2010). "Word retrieval failures in old age: The relationship between structure and function." Journal of Cognitive Neuroscience **22**(7): 1530-1540
Shallice, T. and E. K. Warington (1980). Single and multiple component central dyslexia syndrome. Deep Dyslexia. M. Coltheart, K. E. Patterson and J. C. Marshall. London, Routledge and Kegan Paul**:** 109-145.
Sharma, N., V. M. Pomeroy, et al. (2006). "Motor imagery: A backdoor to the motor system after stroke?" Stroke **37**(7): 1941-1952.
Sharma, N., L. H. Simmons, et al. (2009). "Motor Imagery After Subcortical Stroke A Functional Magnetic Resonance Imaging Study." stroke **40**(4): 1315-1324.
Shergill, S. S., M. J. Brammer, et al. (2002). "Modulation of activity in temporal cortex during generation of inner speech." Human Brain Mapping **16**: 219-227.
Shergill, S. S., D. K. Tracy, et al. (2006). "Timing of covert articulation: An fMRI study." Neuropsychologia **44**: 2573-2577.

Shimony, J. S., A. Z. Snyder, et al. (2004). "The study of neural connectivity using diffusion tensor tracking." Cortex **40**(1): 213-215.
Shuster, L. I. and S. K. Lemieux (2005). "An fMRI investigation of covertly and overtly produced mono- and multisyllabic words." Brain and Language **93**: 20-31.
Silveri, M. C., M. G. Leggio, et al. (1994). "The cerebellum contributes to linguistic production - A case of agrammatic speech following a right cerebellar lesion." Neurology **44**(11): 2047-2050.
Simmons-Mackie, N., T. T. Threats, et al. (2005). "Outcome assessment in aphasia: A survey." Journal of Communication Disorders **38**(1): 1-27.
Smania, N., F. Bazoli, et al. (1997). "Visuomotor imagery and rehabilitation of neglect." Archives of Physical Medicine and Rehabilitation **78**(4): 430-436.
Snowden, J. S. and D. Neary (2003). "Progressive anomia with preserved oral spelling and automatic speech." Neurocase **9**(1): 27-43.
Sokolov, A. N. (1972). Inner Speech and Thought. New York, London, Plenum Press.
Solodkin, A., P. Hlustik, et al. (2004). "Fine modulation in network activation during motor execution and motor imagery." Cerebral Cortex **14**(11): 1246-1255.
Spencer, K. A., P. J. Doyle, et al. (2000). "Examining the facilitative effects of rhyme in a patient with output lexicon damage." Aphasiology **14**(5-6): 567-584.
Staempfli, P., C. Reischauer, et al. (2008). "Combining fMRI and DTI: A framework for exploring the limits of fMRI-guide DTI fiber tracking and for verifying DTI-based fiber tractography results." Neuroimage **39**(1): 119-126.
Stebbins, G. T., D. L. Nyenhuis, et al. (2008). "Gray matter atrophy in patients with ischemic stroke with cognitive impairment." Stroke **39**(3): 785-793.
Stein, D. G., S. Brailowsky, et al. (1995). Brain Repair. Oxford, Oxford University Press.
Stonnington, C. M., G. Tan, et al. (2008). "Interpreting scan data acquired from multiple scanners: A study with Alzheimer's disease." Neuroimage **39**(3): 1180-1185.
Strick, P. L., R. P. Dum, et al. (2009). "Cerebellum and nonmotor function." Annual Review of Neuroscience **32**: 413-434.
Swinburn, K., G. Porter, et al. (2004). Comprehensive Aphasia Test. East Sussex, UK, Psychology Press.
Thiel, A., B. Habedank, et al. (2006). "From the left to the right: How the brain compensates progressive loss of language function." Brain and Language **98**(1): 57-65.
Thuy, D. H. D., K. Matsuo, et al. (2004). "Implicit and explicit processing of kanji and kana words and non-words studied with fMRI." Neuroimage **23**(3): 878-889.
Tyler, L. K., M. A. Shafto, et al. (2010). "Preserving Syntactic Processing across the Adult Life Span: The Modulation of the Frontotemporal Language System in the Context of Age-Related Atrophy." Cerebral Cortex **20**(2): 352-364.
Upton, C. and E. Upton (2004). Oxford Rhyming Dictionary. Oxford, UK, Oxford University Press.
Usui, K., A. Ikeda, et al. (2005). "Processing of Japanese morphogram and syllabogram in the left basal temporal area: electrical cortical stimulation studies." Brain Research. Cognitive Brain Research **24**(2): 274-283.
Vallar, G. and A. D. Baddeley (1984). "Phonological Short-Term Store, Phonological Processing and Sentence Comprehension - A Neuropsychological Case-Study." Cognitive Neuropsychology **1**(2): 121-141.
Vernooij, M. W., M. Smits, et al. (2007). "Fiber density asymmetry of the arcuate fasciculus in relation to functional hemispheric language lateralization in both right- and left-handed healthy subjects: A combined fMRI and DTI study." Neuroimage **35**(3): 1064-1076.
Vigliocco, G. and R. J. Hartsuiker (2002). "The interplay of meaning, sound, and syntax in sentence production." Psychological Bulletin **128**(3): 442-472.
Vigneau, M., V. Beaucousin, et al. (2006). "Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing." Neuroimage **30**(4): 1414-1432.
Vitali, P., J. Abutalebi, et al. (2007). "Training-induced brain remapping in chronic aphasia: A pilot study." Neurorehabilitation and Neural Repair **21**(2): 152-160.

Vul, E., C. Harris, et al. (2009). "Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition." Perspectives on Psychological Science **4**(3): 274-290.
Vygotsky, L. S. (1962). Thought and Language. Cambridge, Massachusetts M.I.T. Press.
Wada, J. and T. Rasmussen (1960). "Intracarotid injection of sodium amytal for the lateralization of cerebral speech dominance - Experimental and clinical observations." Journal of Neurosurgery **17**(2): 266-282.
Wagner, A. D., W. Koutstaal, et al. (2000). "Task-specific repetition priming in left inferior prefrontal cortex." Cerebral Cortex **10**(12): 1176-1184.
Wagner, A. D., E. J. Pare-Blagoev, et al. (2001). "Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval." Neuron **31**(2): 329-338.
Warburton, E., C. J. Price, et al. (1999). "Mechanisms of recovery from aphasia: Evidence from positron emission tomography studies." Journal of Neurology Neurosurgery and Psychiatry **66**(2): 155-161.
Warlow, C. P., M. S. Dennis, et al. (1996). Stroke: A Practical Guide to Management. Oxford, Blackwell Sciences, Ltd. .
Warren, J. E., J. Crinion, et al. (2009). "Anterior temporal lobe connectivity correlates with functional outcome after aphasic stroke." Brain **132**(12): 3428-3442.
Warrington, E. K. and P. McKenna (1983). The Graded Naming Test. London, NFER.
Warrington, E. K. and T. Shallice (1979). "Semantic access dyslexia." Brain **102**(Mar): 43-63.
Wechsler, D. (1955). Wechsler Adult Intelligence Scale. Baltimore Williams & Wilkins.
Weiller, C., F. Chollet, et al. (1992). "Functional reorganization of the brain in recovery from striatocapsular infarction in man." Annals of Neurology **31**(5): 463-472.
Weiller, C., C. Isensee, et al. (1995). "Recovery from Wernicke's aphasia - A positron emission tomographic study." Annals of Neurology **37**(6): 723-732.
Weiller, C., S. C. Ramsay, et al. (1993). "Individual patterns of functional reorganization in the human cerebral-cortex after capsular infarction." Annals of Neurology **33**(2): 181-189.
Weisenburg., T. H. and K. E. Mcbride (1935). Aphasia - a Clinical and Psychological Study. New York, The Commonwealth Fund.
Wernicke, C. (1874). Der Aphasische Symptomenkomplex. Breslau Cohn and Weigert.
West, R. L. (1996). "An application of prefrontal cortex function theory to cognitive aging." Psychological Bulletin **120**(2): 272-292.
White-Thomson, M. (1999). "Naming therapy for an aphasic person with fluent empty speech." The Aphasia Therapy File **1**: 93-100.
White, K. K., L. Abrams, et al. (2008). "Why did I right that? Factors that influence the production of homophone substitution errors." Quarterly Journal of Experimental Psychology **61**(7): 977-985.
Whitworth, A., J. Webster, et al. (2005). A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia, A Clinician's Guide. Hove and New York, Psychology Press.
Wierenga, C. E., M. Benjamin, et al. (2008). "Age-related changes in word retrieval: Role of bilateral frontal and subcortical networks." Neurobiology of Aging **29**(3): 436-451.
Wierenga, C. E., L. M. Maher, et al. (2006). "Neural substrates of syntactic mapping treatment: An fMRI study of two cases." Journal of the International Neuropsychological Society **12**(1): 132-146.
Willinsky, R. A., S. M. Taylor, et al. (2003). "Neurologic complications of cerebral angiography: Prospective analysis of 2,899 procedures and review of the literature." Radiology **227**(2): 522-528.
Wilshire, C. E. (2008). "Cognitive neuropsychological approaches to word production in aphasia: Beyond boxes and arrows." Aphasiology **22**(10): 1019-1053.
Wilson, S. M. and A. P. Saygin (2004). "Grammaticality judgment in aphasia: Deficits are not specific to syntactic structures, aphasic syndromes, or lesion sites." Journal of Cognitive Neuroscience **16**(2): 238-252.

Wilson, S. M., A. P. Saygin, et al. (2004). "Listening to speech activates motor areas involved in speech production." Nature Neuroscience **7**(7): 701-702.

Wingfield, A. and M. Grossman (2006). "Language and the aging brain: Patterns of neural compensation revealed by functional brain imaging." Journal of Neurophysiology **96**(6): 2830-2839.

Winhuisen, L., A. Thiel, et al. (2005). "Role of the contralateral inferior frontal gyrus in recovery of language function in poststroke aphasia: A combined repetitive transcranial magnetic stimulation and positron emission tomography study." Stroke **36**(8): 1759-1763.

Wise, R. J. S. (2003). "Language systems in normal and aphasic human subjects: Functional imaging studies and inferences from animal studies." British Medical Bulletin **65**: 95-119.

Wise, R. J. S., J. Greene, et al. (1999). "Brain regions involved in articulation." Lancet **353**(9158): 1057-1061.

Wise, R. J. S., S. K. Scott, et al. (2001). "Separate neural subsystems within 'Wernicke's area'." Brain **124**: 83-95.

Wisenburn, B. and K. Mahoney (2009). "A meta-analysis of word-finding treatments for aphasia." Aphasiology **23**(11): 1338–1352.

Wright, H. H., R. A. Downey, et al. (2007). Processing distinct linguistic information types in working memory in aphasia. 36th Clinical Aphasiology Conference, Ghent, Belgium, Psychology Press.

Xiang, H. D., H. M. Fonteijn, et al. (2010). "Topographical functional connectivity pattern in the perisylvian language networks." Cerebral Cortex **20**(3): 549-560.

Xu, X. J., M. M. Zhang, et al. (2004). "Cortical language activation in aphasia: A functional MRI study." Chinese Medical Journal **117**(7): 1011-1016.

Yamadori, A., Y. Osumi, et al. (1977). "Preservation of singing in Broca's aphasia." Journal of Neurology Neurosurgery and Psychiatry **40**(3): 221-4.

Yamamoto, S., M. Takasawa, et al. (2007). "Deterioration of hemiparesis after recurrent stroke in the unaffected hemisphere: Three further cases with possible interpretation." Cerebrovascular Diseases **23**(1): 35-39.

Zangwill, O. (1947). "Psychological aspects of rehabilitation in cases of brain injury." British Journal of Psychology **37**: 60-69.

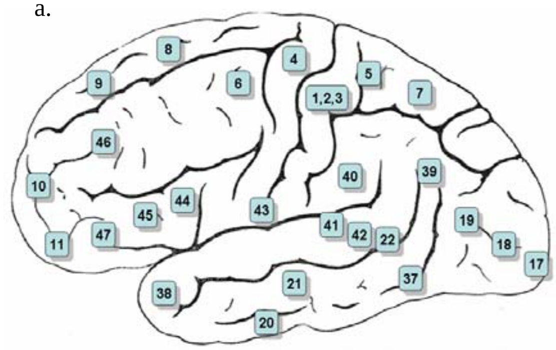
Zanini, S., K. Bryan, et al. (2005). "The effects of age and education on pragmatic features of verbal communication: Evidence from the Italian version of the Right Hemisphere Language Battery (I-RHLB)." Aphasiology **19**(12): 1107-1133.

Zettin, M., S. F. Cappa, et al. (1997). "Agrammatic speech production after a right cerebellar haemorrhage." Neurocase **3**(5): 375-380.

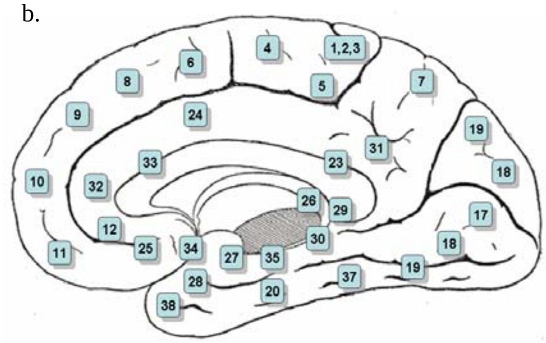
Appendix 1: Neuroanatomical Nomenclature

Brain anatomy, including Brodmman's Areas (BA) on the surface (a) and the midline (b) of the cortex. Anatomical labels on the surface of the cortex (c) (adapted from Wikipedia).

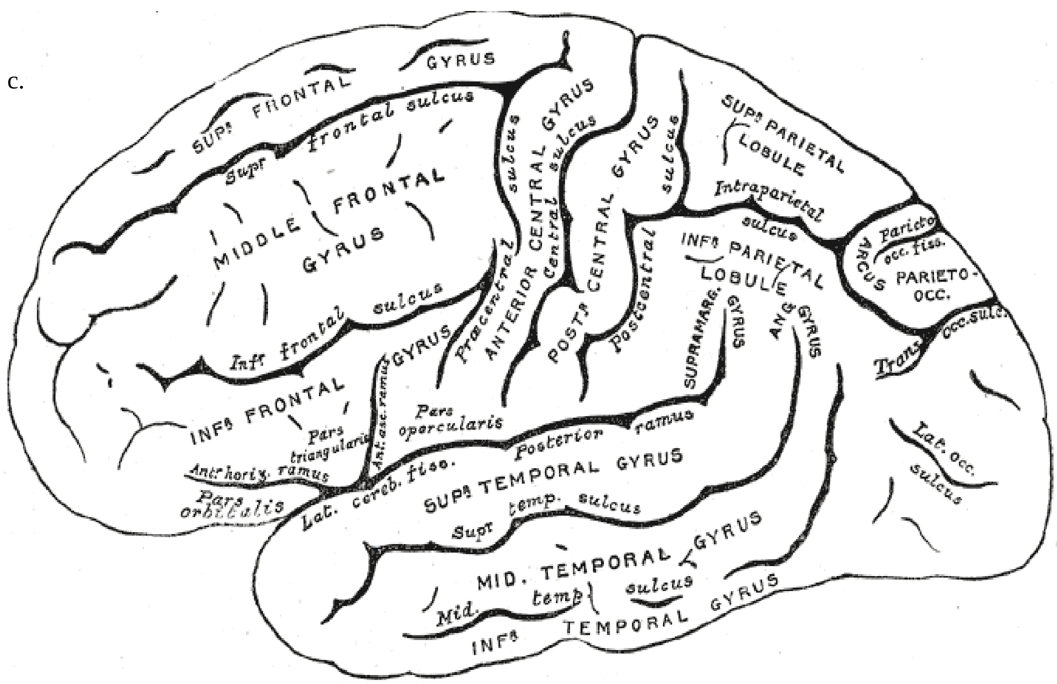
a.



b.



c.



Appendix 2: Patients' Exclusion and Inclusion

This appendix details the inclusion and exclusion of patients in the different parts of the study.

Figure 1. Assessment and testing of the different populations of patients in different parts of the study.

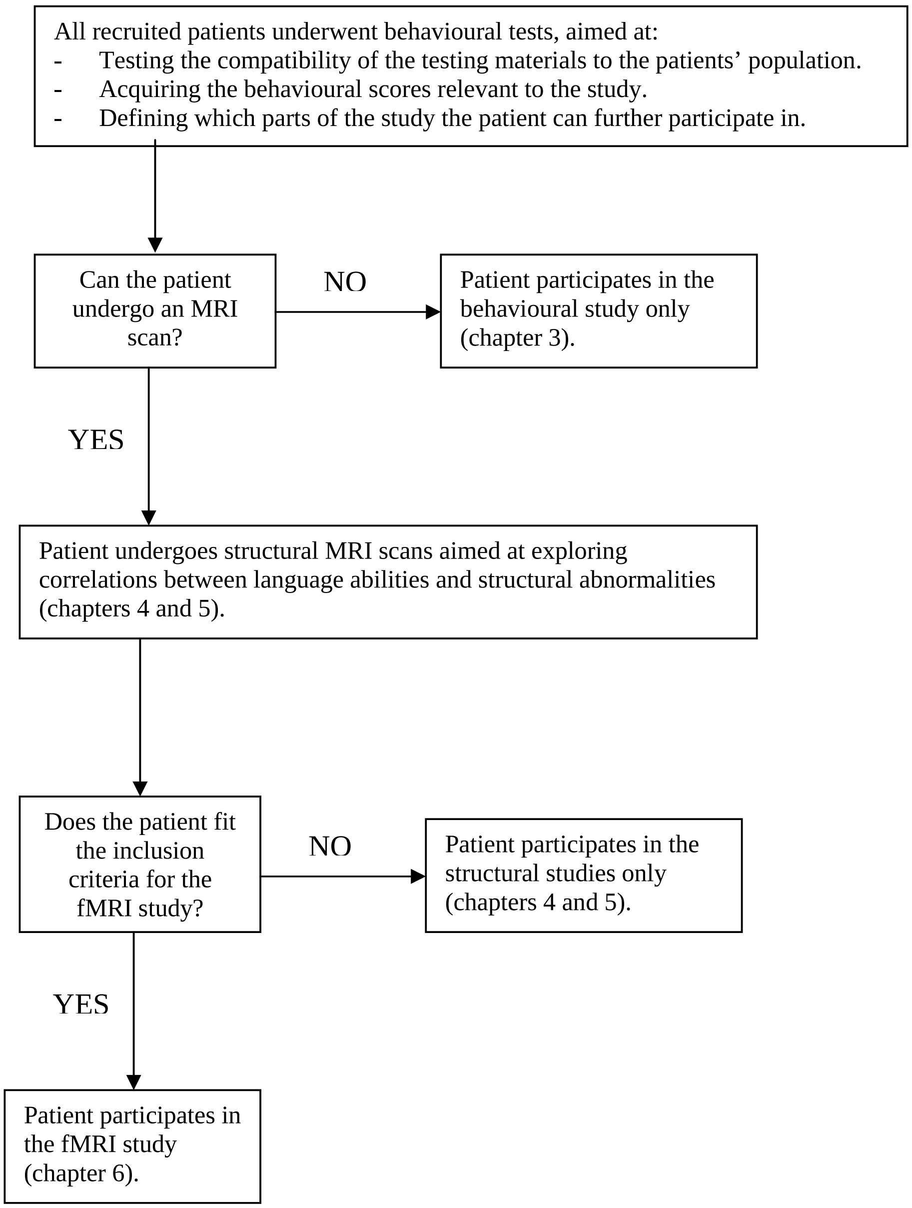


Table 1. Patients' demographic and research related information.

	Patient	Age	Sex	Hand[1]	Type of Stroke	Studies		Time between last stroke and cognitive-behavioural assessment (months)	Time between last cognitive-behavioural assessment and head scan (months:days)	Scan	Lesion volume (ml)
						Structural studies (chapters 4-5)	fMRI study (chapter 6)				
1	AD	66	m	R	ischaemic	+	+	11	2:4	3T MRI	49.904
2	AE	69	m	R	ischaemic	+		22	0:27	3T MRI	198.544
3	AP	73	m	L	ischaemic	+		18	0:26	3T MRI	68.729
4	AS	48	f	R	ischaemic			10	-2:0	CT	
5	BA	62	m	R	haemorrhagic	+	+	10	8:1	3T MRI	9.949
6	BBS	78	m	A (-0.1)	ischaemic	+		64	0:17	3T MRI	234.424
7	CB	69	m	R	ischaemic	+	+	25	0:7	3T MRI	8.762
8	CR	60	m	R	ischaemic			14		none	
9	DB	78	f	R	ischaemic	+	+	9	8:17	3T MRI	0.856
10	DH	78	m	L	ischaemic	+		20	1:10	3T MRI	124.681
11	DR	73	m	R	ischaemic			10	2:10	1.5T MRI	
12	EB	21	f	R	ischaemic	+		15	2:22	1.5T MRI	121.768
13	GH	42	f	R	ischaemic	+		13	4:14	3T MRI	101.806
14	HD	81	m	R	ischaemic	+		19	0:12	3T MRI	129.478
15	IB	62	m	R	ischaemic	+	+	16	9:3	3T MRI	127.46
16	IC	61	m	R	ischaemic			22	-0:24	CT	
17	IF	78	m	R	ischaemic			19		none	
18	JFI	65	f	R	haemorrhagic	+		24	12:27	1.5T MRI	92.106
19	JFO	71	m	R	ischaemic	+		59	2:19	1.5T MRI	304.253

	Patient	Age	Sex	Hand[1]	Type of Stroke	Studies: Structural studies (chapters 4-5)	Studies: fMRI study (chapter 6)	Time between last stroke and cognitive-behavioural assessment (months)	Time between last cognitive-behavioural assessment and head scan (months:days)	Scan	Lesion volume (ml)
20	JHH	71	M	R	ischaemic			111		none	
21	JS	79	m	L	ischaemic	+		8	0:22	3T MRI	21.527
22	LM	49	f	R	ischaemic	+	+	20	1:20	3T MRI	4.679
23	MM	70	m	A (0.4)	ischaemic			29		none	
24	NET	70	m	R	ischaemic	+		87	0:20	3T MRI	224.092
25	PT	65	m	R	ischaemic			23		none	
26	RB	53	f	R	ischaemic	+		36	2:20	3T MRI	201.277
27	RC	55	m	R	ischaemic	+		48	0:29	3T MRI	197.171
28	TH	75	f	L	haemorrhagic			66	1:20	CT	
29	WS	51	F	R	ischaemic	+		24	0:14	3T MRI	161.097

[1] L=Left, R=Right, A=Ambidextrous. In brackets: the score achieved on the Edinburgh Handedness Inventory (1971) for ambidextrous subjects, where (-1)=strongly left handed, (1)=strongly right handed and (0)=completely ambidextrous.

Appendix 3: Behavioural Questionnaires

The lexical stress questionnaire was used in chapter 2 only. Lists 1 and 2 of the rhyme and homophone judgement were used in chapters 2-5. List 3 of the non-word homophone judgement was used in chapter 2.

Lexical Stress

University
Complexity
Apologize
Information
Escalator
Tomatoes
Hotel
Mountain
Slower
Plastic
September
Teacher
Indeed
Cinema
Later
Banana
Pencil
Apartment
Flower
Orange
Happiness
History
Perhaps
Important
Economic
Below
Volunteer
Refugee
Certificate
Difficult
Electricity
Elevator
Elephant
Decide
Under
Photograph
Hurricane
Aluminium
Shortly
Pedestrian

Canada
China
Politician
Photography
Mistake
Computer
Referee
Entertain
Committee
Understand
Before
Contradict
Notebook
Statue
Table
Demand
Operation
Comprehend
Romantic
Correct
Conversation
Supermarket
Psychology
Photographic
Until
About
Above
Satellite
Monday
Acrobat
Palace
Biology
Delicate
Expensive
Activity
February
Photographer
Goodbye

Homophones – List 1		Homophones – List 2	
Non-Words		**Non-Words**	
wime	waim	byme	bime
fick	phic	quib	kwib
bick	blic	zoar	zure
coim	koym	peam	pame
voar	vore	scad	skad
zoal	zole	thib	shib
hain	hine	grex	geks
phex	ffeks	heem	heam
shad	chad	tain	tane
noal	nool	foym	fyme
Words		**Words**	
bury	berry	prey	pray
sea	see	pail	pale
frey	fey	pout	port
route	root	peach	poach
earn	urn	some	sum
kill	sill	maid	made
fury	ferry	pear	pair
beach	beech	new	no
pour	pore	sew	so
ear	oar	dough	doe
might	mite	dual	jewel
sail	soil	shoot	soot
row	rough	neigh	nigh
raid	ride	pea	pie
weigh	way	break	brake
flay	flee	dear	dare
weak	wake	duet	cruet
cell	sell	sore	saw
bore	bow	sight	sigh
quay	key	home	hum

Non-Words Homophones – List 3

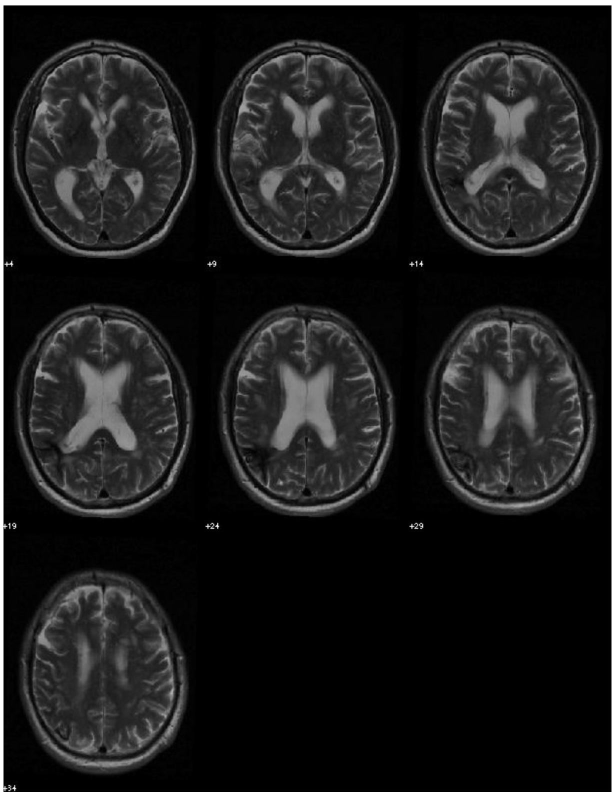
kar	qarr
barious	barius
zait	zeight
bergog	bergogue
maccep	maxep
ama	amah
boarnaf	borenaf
fip	phif
thirm	theerm
celb	selb
florc	phlork
modate	nodait
galp	gelp
lail	lale
pight	pyte
chlott	chot
wut	woot
ghuff	juff
meer	mier
plask	plasque
clarp	klarb
yarma	yarmah
tousen	towsen
chirr	shirr
acalian	akkalian
amel	amle
shink	shynke
sarg	tharg
doal	lole
crid	croad
mol	mil
vrack	traque
hine	hain
qaffal	kaphal
cibe	kighb
tade	taide
sert	soort
twam	twim
pailey	pailie
klid	klide

Rhymes – List 1		Rhymes – List 2	
town	gown	pleat	treat
rush	gush	ghost	roast
boot	flute	creak	break
pea	play	paw	hour
sort	part	match	hatch
wart	fort	hive	dive
hush	bush	bowl	mole
sauce	worse	flair	year
bone	cone	tone	gone
bond	hand	zoo	thou
tint	pint	low	toe
tweak	freak	down	flown
dull	hull	comb	gnome
call	ball	rose	lose
pool	wool	mint	hint
four	saw	bait	skate
wand	pond	horse	force
jute	foot	glove	wove
food	blood	card	ward
bear	chair	gull	full
boast	cost	fall	shall
batch	watch	doe	cow
shoe	screw	love	dove
cheat	sweat	date	plait
head	bed	sea	quay
chew	hoe	wed	bead
five	give	dome	bomb
fool	tool	yard	hard
pose	prose	you	two
hole	owl	mood	brood

Appendix 4: Patients' Case Histories

Below are the descriptions of the patients who participated in the study. All patients are native speakers of British English, monolinguals and right handed, unless stated otherwise. The axial slices showing the lesion are presented for each patient's T2 scan (or CT scan when MRI is not available). Left side of the brain is on the left side of the image.

BA

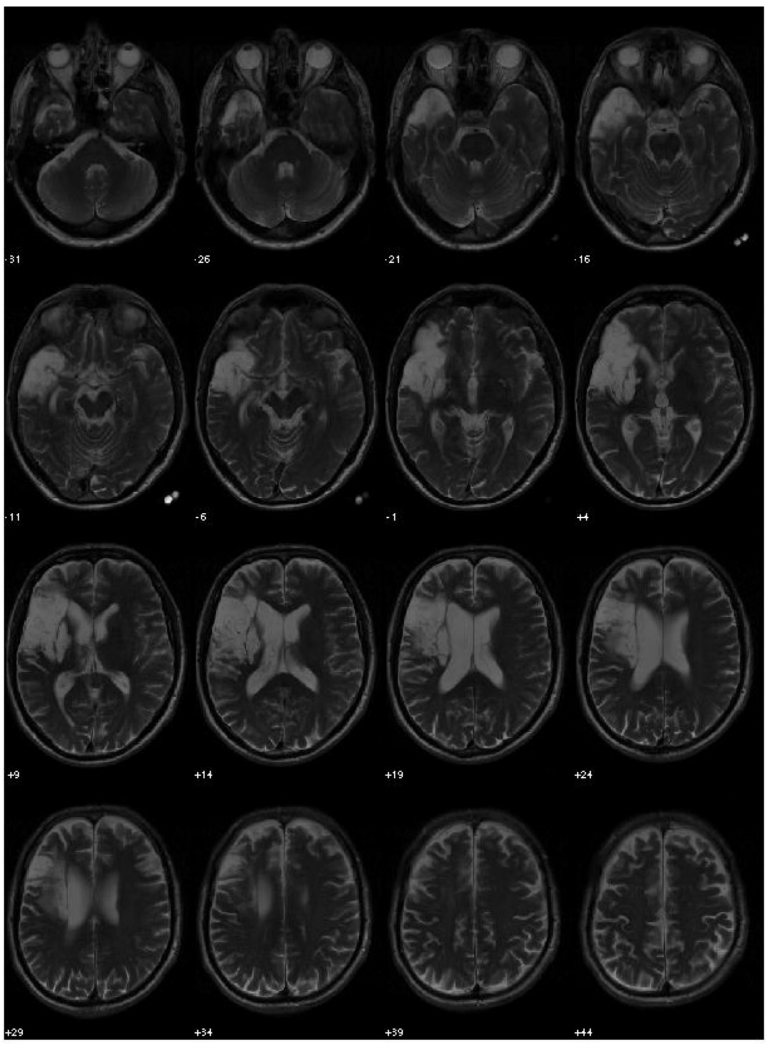


BA is a chief executive of a charity organisation, who had a left parietal haemorrhagic stroke while driving his car, at the age of 61. The cause of the stroke was hypertension which was subsequently successfully controlled by medication. On admission to the hospital he presented with mild semantic impairment and word finding difficulty for low frequency words. On semantic fluency his performance was very low. He also had impairment in arithmetic, memory and frontal-executive functions, as assessed by a neuropsychologist. An acute CT performed within 24 hours showed a left posterior parietal intracranial haemorrhage, extending into the posterior horn of the lateral ventricle. Angiogram was normal and there was no evidence of stenosis or arterio-venous malformation (AVM). He received inpatient rehabilitation.

When tested 10 months after the stroke he had completely recovered, performing at ceiling on all tests (comprehension, production, repetition, inner speech and cognitive tests). He had no semantic, phonological or grammatical errors. He reports that occasionally he still has mild word finding difficulties. MRI scan performed 18 months post stroke showed a volume loss in the angular gyrus of the left inferior parietal lobe. FLAIR imaging showed multiple white matter lesions in the deep and periventricular white matter regions.

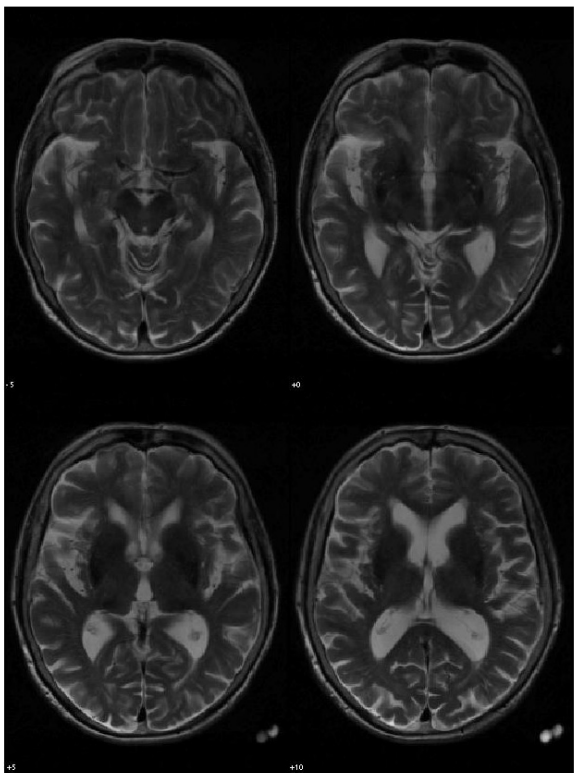
IB

IB is a computer engineer who had a left MCA ischaemic stroke at the age of 60, due to Atrial Fibrillation (AF). On admission he presented with right sided weakness, facial droop, loss of speech and inability to follow commands. An acute CT performed within 24 hours showed reduced signal in the cortical and subcortical white matter of the left fronto-temporal operculum, the left insula and left lentiform nucleus. Carotid Doppler was normal. He received speech and language therapy as an inpatient and in the community.



When tested 16 months after the stroke he had a mild right side hemiplegia and non-fluent, slow and hesitant spontaneous speech but with very few errors. He has occasional word finding difficulty but can easily get a message across. On formal assessment of reading, writing, comprehension and production his performance was at ceiling level. His difficulties can be seen mainly in the oral and written picture description tasks. IB's difficulties are evident by the time it takes him to complete tasks, rather than by the error rate: IB works very slowly and takes a long time to solve relatively easy language tasks. His inner speech is well preserved, and his performance on the homophone and rhyme judgement was almost at ceiling. He noticed that inner speech tasks were more difficult for him than overt speech tasks. When measuring the time it took him to complete the tasks, it was found that while rhyme judgement using overt speech took 3:28 minutes, the same task using inner speech took almost double the time: 6:44 minutes. An MRI scan performed 25 months post-stroke showed a lesion involving the temporal pole, anterior parts of the middle and superior temporal gyri, extending into the insula, IFG and left putamen. The adjacent lateral ventricle shows dilatation.

DB

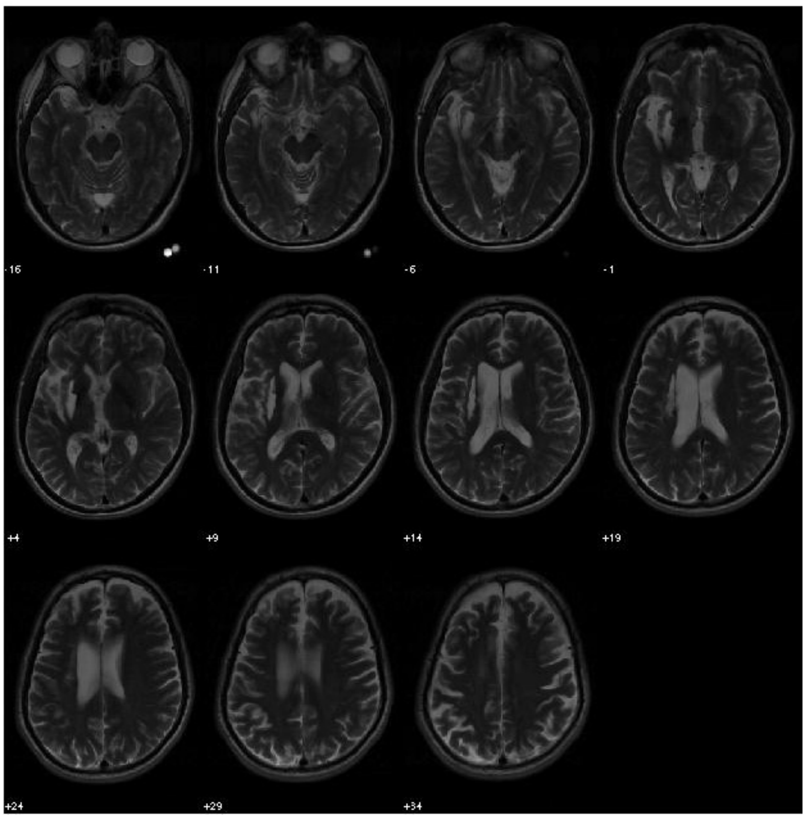


DB is a retired teaching assistant, who had a left MCA ischaemic stroke at the age of 77. The cause of the stroke was AF. Within a week, she had three transient ischaemic attacks (TIAs) in which she could not speak or write and her comprehension was poor. However, she recovered from each episode within 24 hours. An acute CT scan performed within 24 hours showed no abnormality. An MRI scan performed within the same week showed an inferior left insular lesion, most likely caused by small vessel ischaemia. Carotid Doppler was normal. She recovered within a week and received no speech therapy thereafter.

When tested 10 months after her stroke she reported sporadic word finding difficulty, which she described as a sudden ‘black-out’ rather than a tip-of-the-tongue feeling. This was also evident in the phonological fluency test. Infrequently she made semantic errors, both in production and comprehension. In writing she often omits function words, thus creating mildly agrammatic sentences. Her inner speech was well preserved. An MRI scan performed 19 months post-stroke revealed multiple small foci of increased intensity in the white matter of both hemispheres in keeping with small vessel disease, as well as a small left infarct in the insular region.

CB

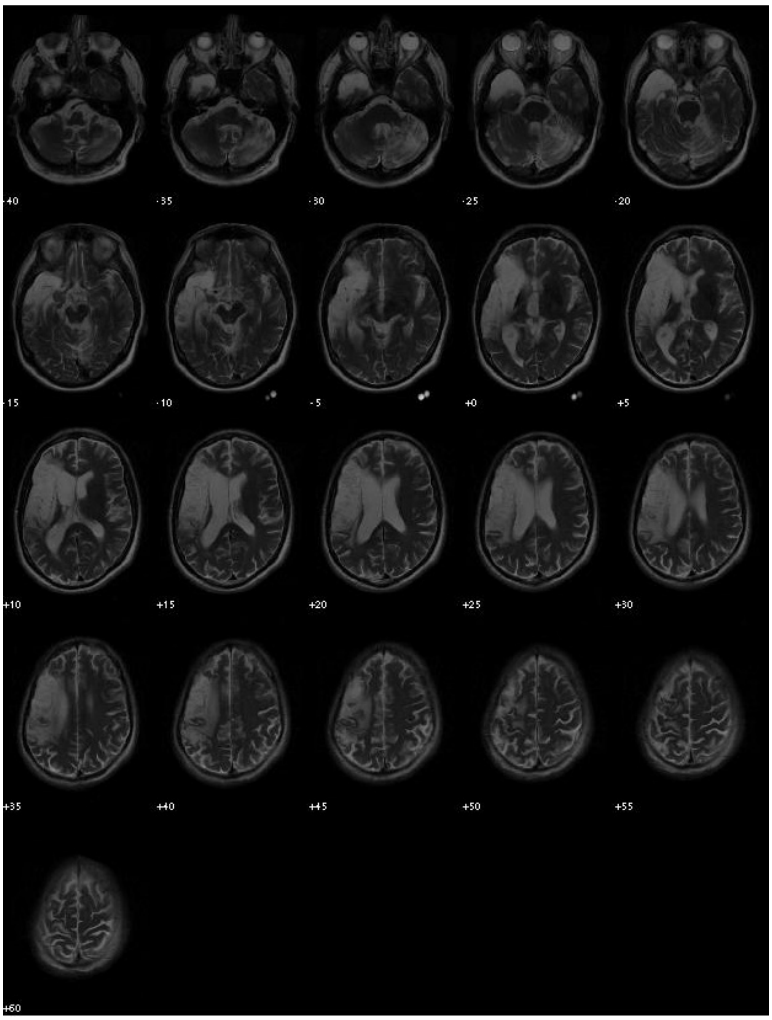
CB had a left MCA ischaemic stroke at the age of 67, due to AF. He worked as a personnel manager in an electric company and is a magician in his free time. Prior to the stroke he had two TIAs, the first occurred 7 years before the stroke but left no visible structural damage (according to a later CT scan), and the second occurring a day before the stroke. He was admitted with right side hemiparesis and severe global aphasia. Initially he could follow two stage commands and gave reliable yes/no answers. However, sentence comprehension was poor and speech production profoundly impaired with severe word finding difficulty, semantic errors, and output further affected by dysarthria. Altogether output was minimal. An acute CT scan performed within 24 hours showed a lesion involving the subinsular region and adjacent basal ganglia. A Carotid Doppler showed significant stenosis. CB received speech and language therapy.



Two years later he is completely mobile and back driving. He performs almost at ceiling on all language tests. Comprehension of written and auditory material is only very slightly impaired as well as repetition, speech production and writing. Fluency is impaired, with only 2 words produced in the phonological fluency task. Semantic fluency is slightly better with 12 words produced in one minute. Lastly, he has no symptoms of apraxia, but has a right side facial weakness. Executive functions, as tested with the Brixton test, are impaired, and visuo-spatial abilities are intact. An MRI scan performed 25 months after the stroke revealed an infarct in the left putamen, insula and subinsular white matter, and low signal intensity in the temporal pole.

BBS

BBS is a retired ambidextrous pilot and engineer. He had a right cerebellar ischaemic stroke at the age of 72, from which he made complete recovery. His second ischaemic stroke occurred two and a half years later, and affected the left MCA territory. It was due to AF. On admission he presented with aphasia, dense right hemiplegia and drowsiness. Two CT scans, the first performed within 24 hours and the second the day after, showed low density in the insular and adjacent frontal and temporal lobes. Carotid Doppler was normal. Initially he was not interested in receiving any speech and language therapy, but started with treatments some years after the stroke.



When tested 5 years and 4 months after the second stroke, at the age of 78, he was left with severe language difficulties and right-sided hemiplegia. Repetition and reading aloud are severely impaired, and show a word length effect. Oral naming and spontaneous speech are also severely affected by both his apraxia and aphasia. His writing is better preserved than his speech; he can often write words which he cannot say aloud. However, he shows very frequent perseverations in writing. For example, in his picture description he wrote the word 'picture' three times. His reading and auditory comprehension are only slightly impaired. The auditory comprehension impairment might be largely attributed to a deficit in short-term verbal memory: BBS asks to repeat sentences very often. He does so by repeating the first part of the sentence and then indicating that he cannot remember the rest of it.

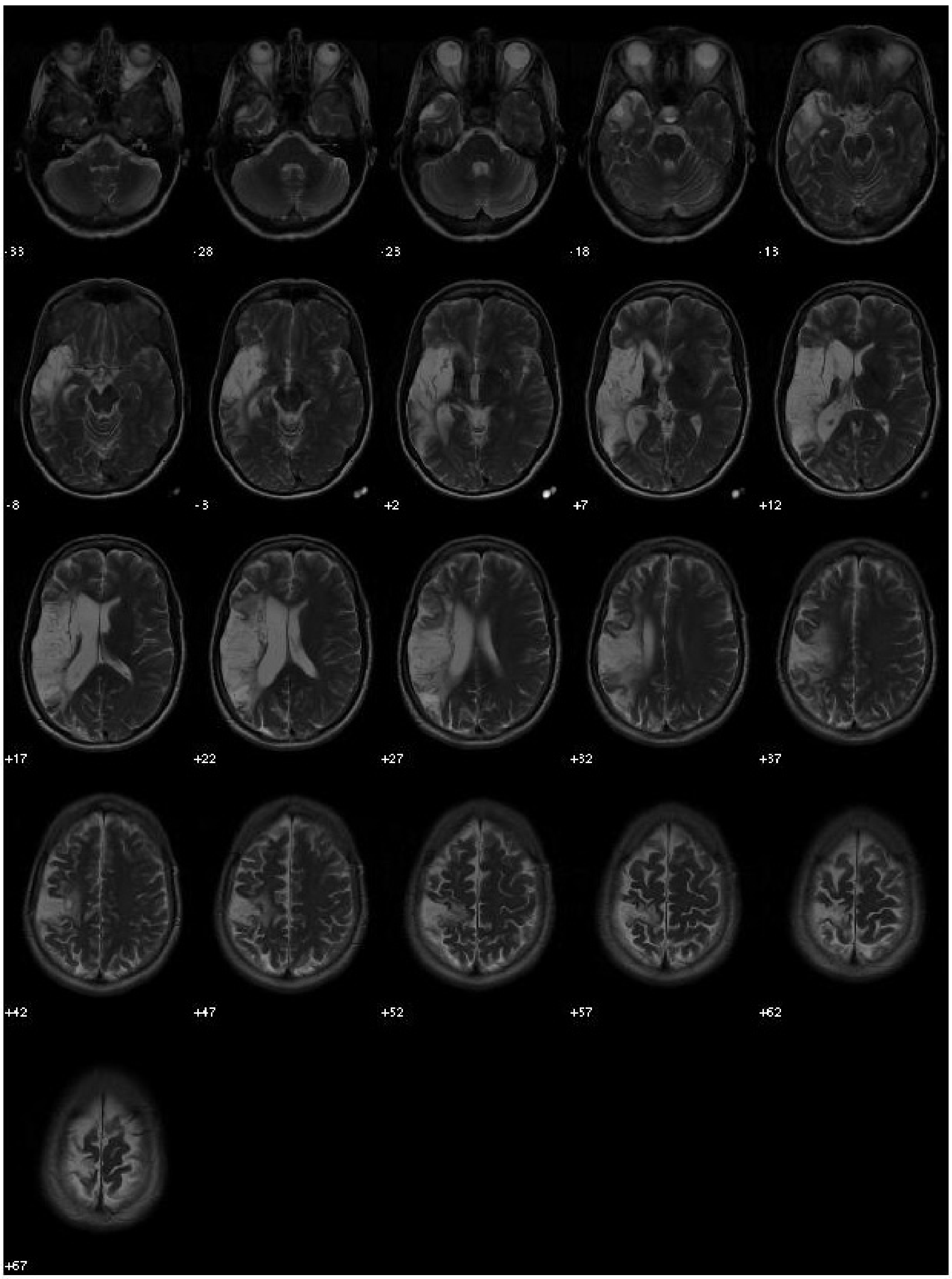
The rhyming and homophone tasks were conducted using mainly inner speech due to his severe apraxia. On the homophone task, he performed well above chance for the words, but not for the non-words. When performing the tasks he sometimes attempts to read aloud, but this usually sounds like a mumble or jargon. Similar results were obtained for the rhyming task. Results of both the computerised and the paper-based tasks were well above chance. BBS attempted a picture rhyming task but this was discontinued after a few trials, the main problem being that due to his severe output it was impossible to establish that BBS is using the right word for each picture. Moreover, since his verbal memory is impaired, he found it difficult to learn the words corresponding to the pictures, before conducting the rhyming task.

An MRI scan performed 5 years and 5 months after his second stroke revealed a lesion involving the temporal pole, anterior and middle parts of the middle and superior temporal gyri, insula, IFG and left putamen, extending into the pre- and post-central gyri.

RB

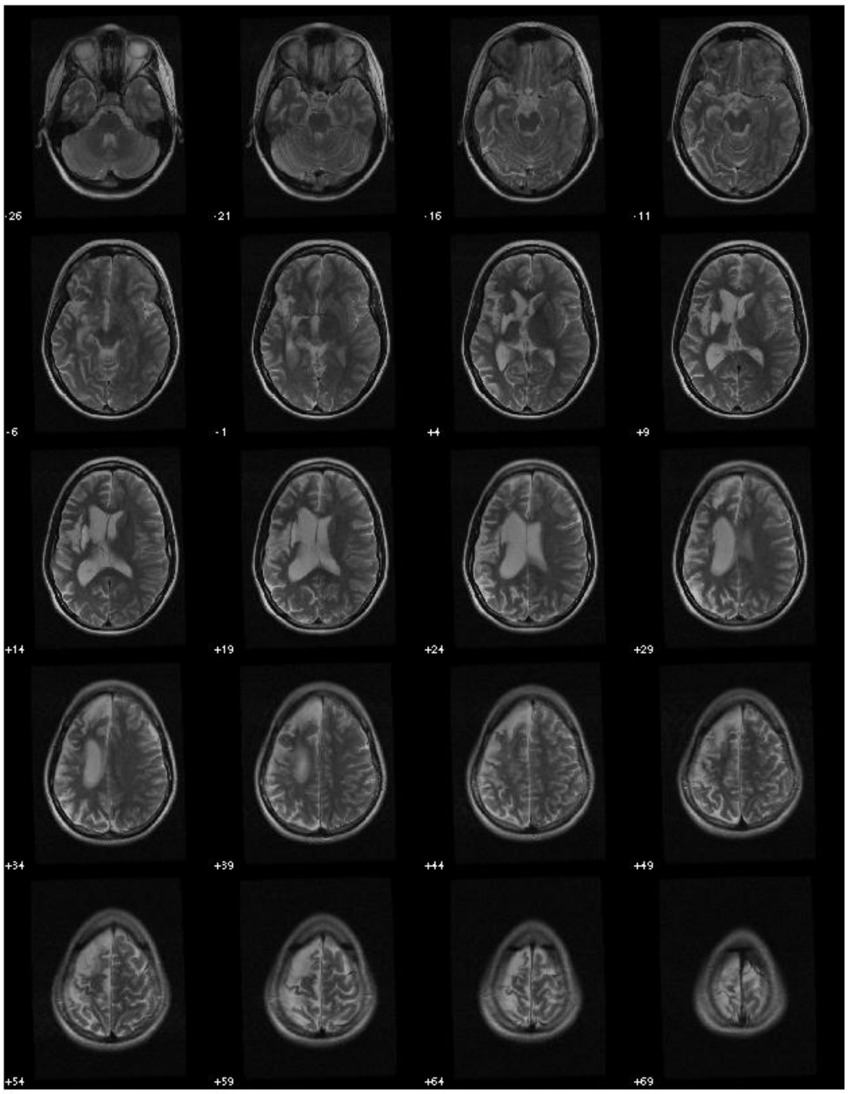
RB used to be a special-needs teacher. She had a left MCA ischaemic stroke at the age of 51. The cause was a deep vein thrombosis (DVT) together with a patent foramen ovale (PFO). It resulted in severe expressive dysphasia, right hemiparesis and motor apraxia, affecting both hand and speech. No acute imaging was available since she has been admitted to a different hospital.

When seen 36 months after the stroke, her right hand was weak and she walked aided by a stick. On examination, semantic knowledge is preserved as well as her memory and non verbal reasoning. Her fluency is markedly impaired and characterised by word finding difficulty, affecting all grammatical categories. Her comprehension is completely functional but on formal testing was impaired. Sentence comprehension is better performed written than spoken. Word repetition shows a mild impairment as well. Reading aloud is severely impaired, marked by occasional signs of deep dyslexia (e.g. read ‘century’ as ‘hundred’ and ‘yacht’ as ‘ship’). Her writing is severely impaired as well. Phonological cues help her name objects and read aloud words in almost every trial. However, she requires a relatively long cue, usually of at least two syllables. Lastly, she states that she hardly has any inner speech, and formal testing confirmed it, with her performance on the inner speech tasks being at chance level. An MRI scan performed 34 months post-stroke revealed a large volume left MCA infarct, involving the left temporal pole, anterior middle and superior temporal gyri, insula, IFG and left putamen, extending into the posterior part of the superior temporal gyrus and the inferior parietal lobe, and further into the inferior part of the superior parietal lobule.



EB

EB suffered an ischaemic stroke at the age of 19, as a result of a DVT together with a PFO. On admission she had global aphasia, with unreliable yes/no answers, and difficulty following one stage commands. Performance was improved when examiners repeated the question or used written material. An acute CT scan performed within 24 hours showed subtle low attenuation in the in the left lentiform nucleus with loss of grey-white matter differentiation. MRA was normal.



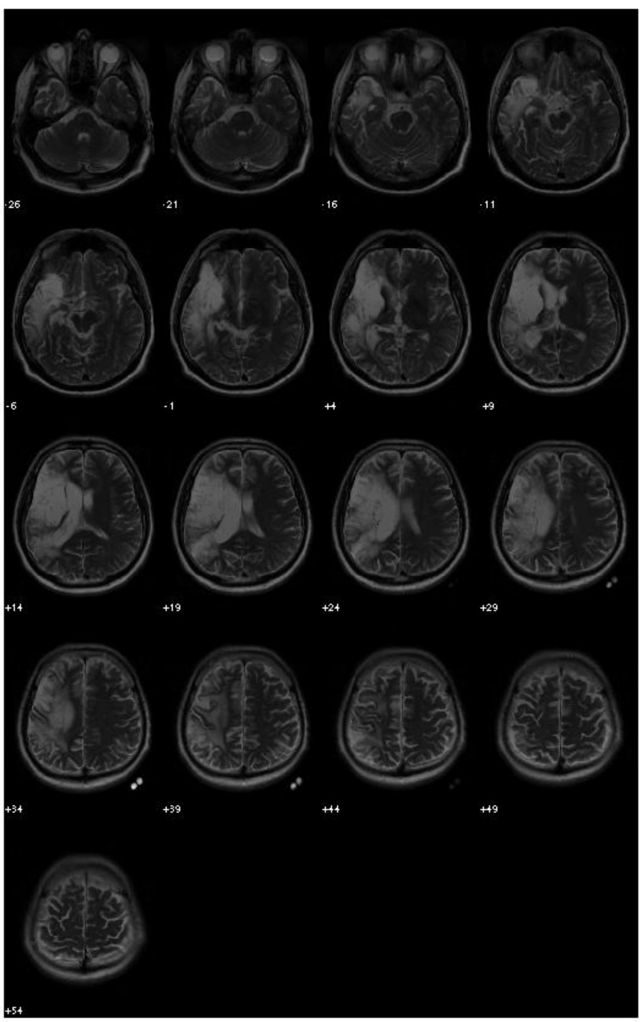
When seen 15 months after the stroke she is left with right side hemiplegia and severe language difficulties. Her right leg is functional but her right hand is completely plegic. She has very limited speech output and mild speech apraxia. Her errors in production, comprehension and reading aloud are mainly semantic. Reading and auditory comprehension are equally mildly impaired but functional. Repetition is limited for long words but preserved for short words of all types. Speech and writing are limited to single words or very short sentences, which are slightly agrammatic.

Reading aloud is impaired and she reports that often she cannot recognise some words. She is unable to read or write non-words. She has no other cognitive impairments.

Due to a patent foramen ovale (PFO) closure device procedure she could not be scanned with a 3T scanner, and instead was scanned in the hospital with a 1.5T MRI scanner. This scan, performed 17 months after the stroke, showed three foci of lesions in the left MCA and ACA regions: the left insula and putamen, extending superiorly into the pre-central gyrus, anterior middle/superior frontal gyrus, and inferior parietal lobe.

RC

RC worked as a maintenance engineer in a big pharmaceutical company. He had a left MCA ischaemic stroke at the age of 51, due to a left carotid dissection. On admission he presented with severe global aphasia, dysphagia and right hemiparesis of both upper and lower limb. An acute CT scan performed within 24 hours revealed a large left MCA infarct. On discharge, three months after the stroke onset, his speech comprehension improved and was almost intact, while speech production was still very limited. He was able to walk aided by a stick but had no movement in the upper limb.



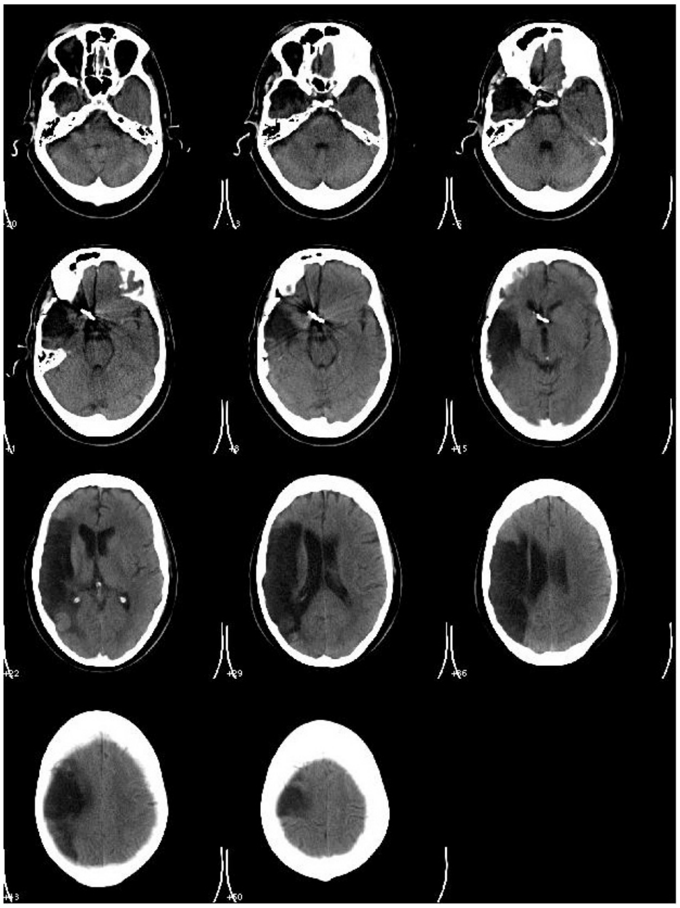
He was formally assessed 4 years after the stroke. He was left with a right hemiplegia, worse in the arm than the leg. He has severe dysphasia affecting both production and comprehension.

His speech is non-fluent, characterised by stereotypical speech (‘Jesus’, ‘OK’, ‘you name it', 'hang about'). He makes semantic errors in speech production. While sometimes noticing the error, he is seldom able to find the correct word. In comprehension tasks, he often confuses semantic distractors, although if given enough time he arrives at the right answer. His auditory comprehension is more preserved than his reading comprehension. He is unable to read or write words in isolation. His reading comprehension is based to a large extent on guess work based on context and implicit knowledge of grammatical structures. He shows no memory or visuo-spatial deficits. MRI scan conducted 4 years and 1 month post-stroke revealed an established infarct in the left hemisphere in the distribution of the MCA, extending from the temporal pole, into the middle and superior temporal lobes, the inferior frontal gyrus, insula and the inferior parietal lobe. In the superior parietal lobe an area of white matter lesion is seen.

IC

IC worked in factory logistics and as a jeweller. He had a left frontal subarachnoid hemorrhage (SAH) at the age of 50, caused by an arteriovenous malformation (AVM). This was treated by aneurism clipping. A second stroke, ischaemic in nature, occurred ten years later, at the age of 60 and a cause was not clearly established. On admission he presented with severe global aphasia, a left homonymous hemianopia, left hemiparesis and severe limb and facial apraxia. CT scans were performed acutely, within 24 hours and after 5 days. Together they showed a very large infarct in the left anterior circulation territory and an associated midline shift to the right of approximately 14 mm. Carotid Doppler was normal.

He was tested 19 months after the second stroke. While his comprehension of spoken language is relatively preserved, his reading comprehension is impaired, especially of long complex sentences. Repetition, reading aloud and language production are very limited. Writing is better preserved but is limited to single words. IC cannot access inner speech from written material, as can be seen with his chance level performance on the homophone and rhyme judgement tasks. His short-term memory is highly impaired, as can bee seen in both the Rey-Osterrieth figure and the CAT recognition memory test.

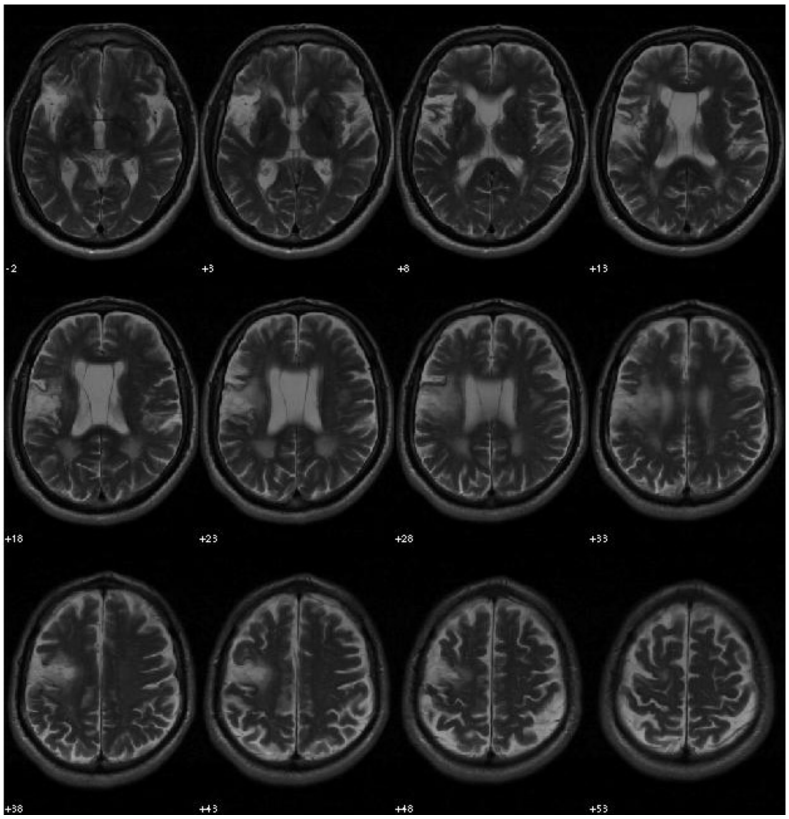


Due to the clip he could not have an MRI scan and therefore had a CT scan. This CT scan was performed one year and 8 months after the second stroke, showing a large established MCA infarct on the left, including the temporal pole, middle and superior temporal lobes, and inferior and superior parietal lobe.

AD

AD is a policeman who suffers from Chronic Inflammatory Demyelinating Polyneuropathy (CIDP). He is a native speaker of American English and had an American accent before his stroke. He has been living in the UK for over 25 years. After being given treatment for CIDP with plasmapheresis at the age of 65, he had a left MCA ischaemic stroke, caused by a left carotid occlusion. He initially presented with language deficits and speech apraxia, as well as a mild right hand weakness. An MRI scan showed multiple small foci of increased intensity in the white matter of both hemispheres in keeping with chronic small vessel disease. In addition, DWI showed increased intensity in the left internal capsule, left frontal and parietal white matter. A Carotid Doppler and an angiogram showed significant stenosis of the internal carotid. This was later treated in a carotid endarterectomy. He received speech and language therapy.

When tested 11 months after the stroke the right side weakness was resolved. Spontaneous speech, reading aloud and repetition are all highly affected by speech dyspraxia. Oral and written comprehensions as well as writing are preserved. He shows very poor performance on the Brixton test but normal performance on all other non-language tests. His inner speech is mostly preserved and despite his abnormal speech output he is able to perform the inner speech tasks quite well.

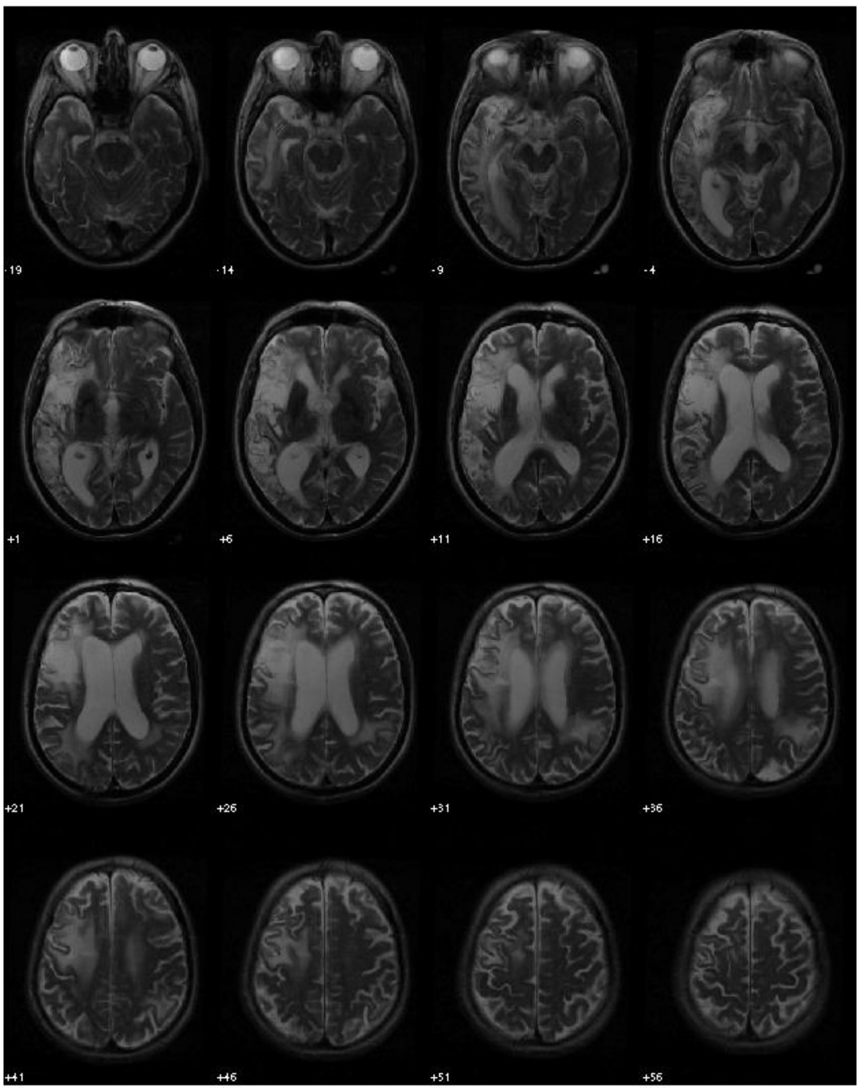


An MRI scan performed 13 months post-stroke showed left insular infarct which extends into the post-central gyrus. Moderate patchy white matter signal change in both cerebral hemispheres can be seen as well, in keeping with chronic small vessel ischaemia.

HD

HD was a marine engineer up to his retirement. He had a first ischaemic stroke at the age of 75, followed by a second ischaemic stroke 5 years later, due to AF. Both involved the left MCA territory. HD lost his speech completely after the first stroke but had made good recovery, supported by speech and language therapy. After the second stroke he presented with fluent aphasia, characterised by speech lacking in semantic content and some comprehension deficits. An acute CT scan performed within 24 hours of his second stroke showed clear evidence of old previous infarcts in the right cerebellar and parietal regions. Carotid Doppler was normal.

He was tested 19 months post-stroke. On formal assessment he shows mild comprehension deficits, although it is clear that his comprehension is functional, as seen in the spoken paragraph comprehension which he completed successfully. A semantic memory task was completed successfully but with many clear hesitations and delays. When debating between answers, he indicated that the semantic distractors confuse him. His errors are purely of the semantic type. Repetition is highly impaired for long words and non-words. HD could not repeat sentences or digit strings. His spontaneous speech is highly fluent and empty, characterised by non-specific expressions like: ‘this fellow’, ‘that thing’. Naming is severely impaired. HD can occasionally retrieve low frequency words or the last part of the word (e.g ‘poon’ for spoon, ‘raffe’ for giraffe, ‘ox’ for socks). Both his reading aloud and writing are severely impaired and he is unable to read non-words at all. He comprehends the meaning of written words but cannot read them aloud. On the homophones and rhyme judgement tasks he performed just above chance. HD attempted the rhyming pictures task, but was unsuccessful because of word finding difficulty. Even when given the word and repeating it a few times, it is forgotten within a couple of minutes. The task was therefore not completed.

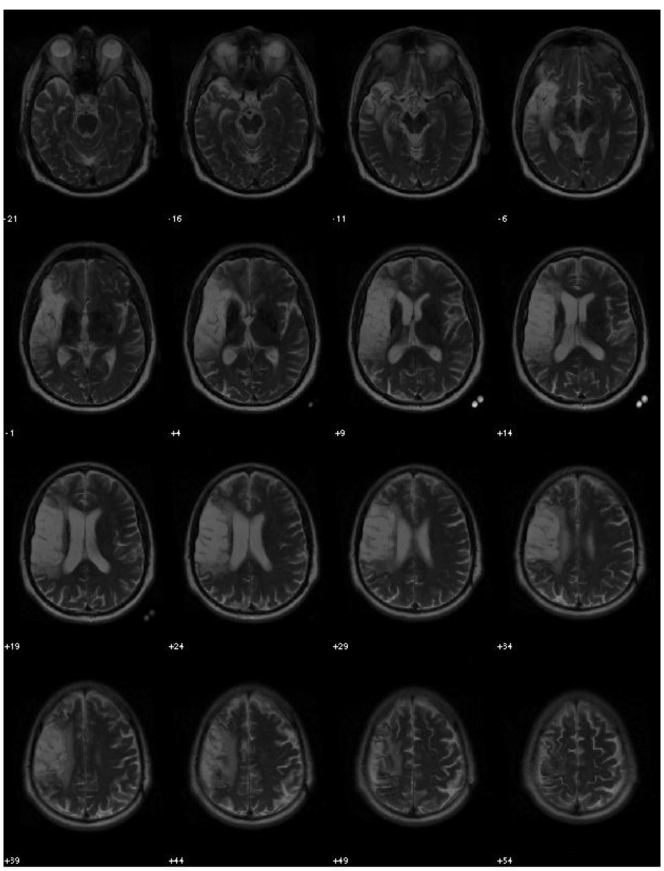


It should be noted that at the time of testing his wife suspected that he is developing a semantic impairment, which she described as gradual and progressive, rather than stable since his last stroke. She described a few instances in which he confused tools and items in the household. For example, she found him digging with his keys in the garden, confusing the washing machine with the dishwasher, and using a teaspoon for eating soup. He is currently under assessment by the Addenbrooke's memory clinic.

An MRI scan performed at the time of cognitive and language assessment, revealed signal change and volume loss in the left inferior and middle frontal lobe, middle and superior temporal lobe, and the inferior parietal lobe. Signal change is also seen in the white matter of the right parietal lobe.

AE

AE is an investment banker who had a left MCA ischaemic stroke at the age of 67. Carotid Doppler showed critical stenosis of the internal carotid, which was later confirmed with MRA. He later went on to have a carotid endarterectomy. Since then he had two stroke related seizures, but imaging did not reveal any new infarcts. On admission to hospital he presented with expressive dysphasia, dyspraxia and right sided hemiparesis. An MRI scan revealed an extensive, mainly cortical, infarct in the territory of the left MCA, partially involving the striatum.



Tested 22 months after the stroke he has a right hand hemiplegia and is walking aided by a stick. He is currently taking seizure medication that makes him easily tired. He is very frustrated about his disability, to which he has good insight.

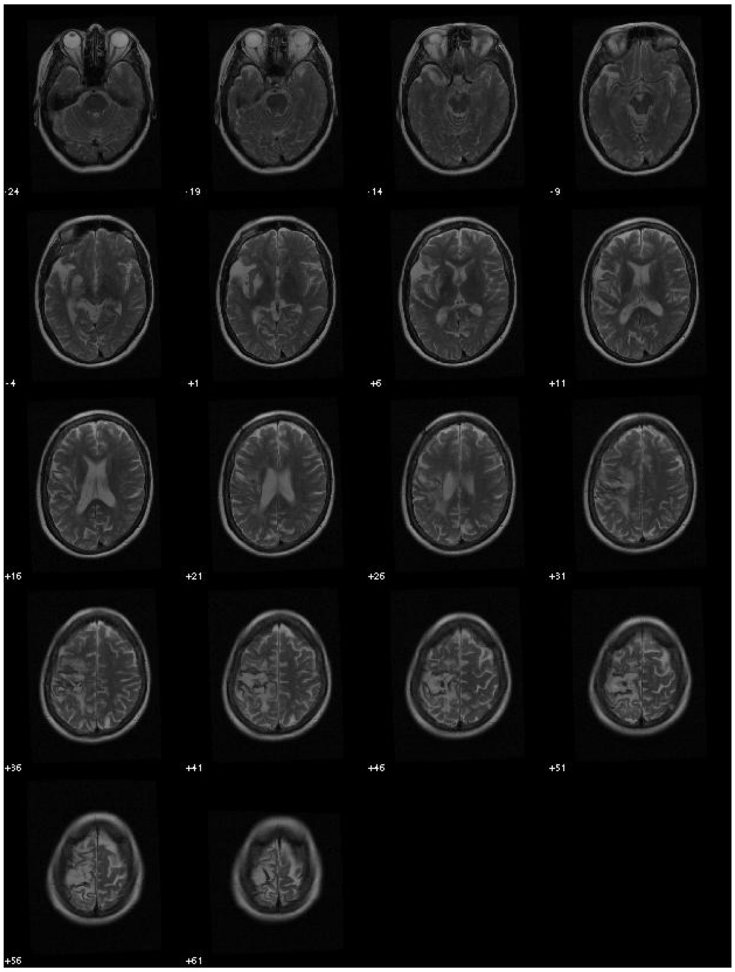
His speech output remains severely limited; he can only say 'yes' and 'no', and cannot read aloud or repeat. He can write single words, but makes many spelling mistakes. He shows word finding difficulty in writing. AE makes semantic errors in comprehension tasks, but seems functional otherwise. His reading comprehension is better than auditory comprehension. AE has both limb and facial apraxia. He is unable to repeat more than one syllable, and shows hesitations and searching behaviour before successful repetition. AE cannot read aloud therefore all inner speech tasks were performed covertly. With an average error rate of 49.2%, it seems that AE has no inner speech, as he himself reports.

An MRI scan was performed at the time of cognitive and language assessment, revealing an established left MCA infarct involving the middle and superior temporal lobes, as well as the parietal lobe.

JFI

JFI had a left hemorrhagic stroke at the age of 63 for which a reason was not established. She worked in the council registry and in the hospital. She was admitted with severe dysphasia and right hemiplegia, and an acute CT scan performed within 24 hours showed a left frontal haematoma. She has no aneurysm and no AVM.

She was tested two years after the stroke. Her verbal output is severely affected. She has specific difficulties with longer words, words containing ‘s’ or ‘r’ sounds and words starting with two consecutive consonants. She also mixes 'p' and 'b'. Typical reading or repetition errors are ‘besident’ for ‘president’, ‘sorn’ for ‘scorn’, ‘gavity’ for ‘gravity’ and ‘tobato’ for ‘tomato’. Language comprehension is mainly preserved, but she has difficulty with complex sentences when presented in written or spoken form. She found the fluency task very difficult. Despite severe language production errors, she performed flawlessly on the inner speech task of homophone judgement, and above chance on the rhyming task.



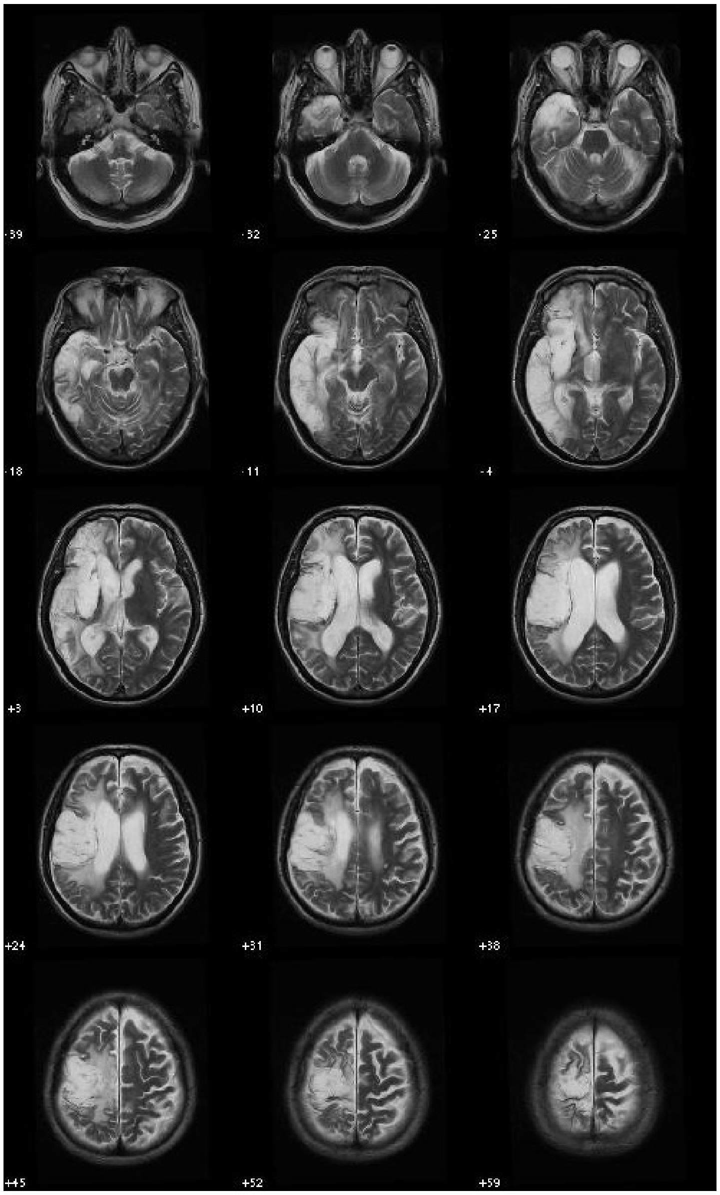
Lastly, on subtest III of the Apraxia Battery for Adults she performed with no difficulty. Whistling was the only task she found difficult but managed to do it independently after a few attempts.

Her short-term memory is severely impaired, as seen from her performance on the CAT recognition test and the Rey-Osterrieth figure test.

Due to having a coronary artery stent she could not be scanned with a 3T MRI, and instead was scanned in the hospital with a 1.5T MRI scanner. This scan was performed a year after the cognitive and language assessment, and revealed lesions in the left putamen, insula, and the adjacent pre- and post-central gyri, extending posteriorly into the superior parietal lobule. A few scattered hyperintensities were visible in the deep white matter of both cerebral hemispheres, consistent with small vessel disease.

JFO

JFO had a left MCA ischaemic stroke at the age of 66, due to AF. Up until his retirement he worked as a builder. On admission he presented with expressive dysphasia, dyspraxia, dense right side hemiparesis and right homonymous hemianopia. He could only follow two stage commands. An acute CT scan performed within 24 hours showed a large infarct in the MCA territory.



He was tested 5 years after the stroke. He is left with right hemiplegia, facial and limb apraxia, memory and acalculia and severe dysphasia. He shows frequent semantic errors in both production and comprehension. His comprehension deficit is more severe in the written domain. He has minimal output, limited to single words, and he can hardly repeat any words. When naming he makes both phonological and semantic errors, and perseverations. He responds to phonological cues in almost 100% of the trials.

He performed just above chance on the inner speech homophone task but at chance level on the rhyming task. Due to a closure procedure for patent foramen ovale (PFO) he could not be scanned with a 3T scanner, and instead was scanned in the hospital with a 1.5T MRI scanner, 2.5 months after performing the language and cognitive assessments. The scan revealed a large left MCA infarct involving the frontal, parietal, temporal and insular regions.

IF

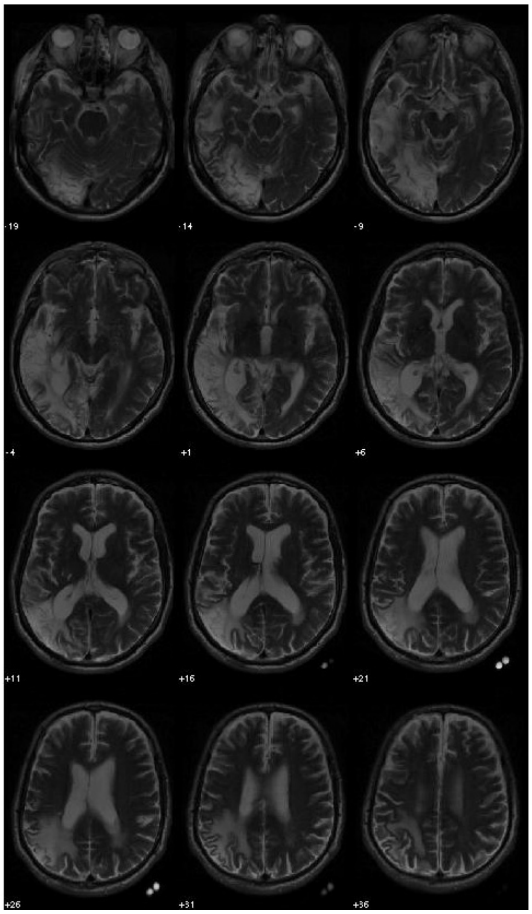
IF, a retired managing director, had two ischaemic strokes at the age of 76. The first stroke left him with a right sided hemiparesis. The second stroke occurred 3 months later and was due to AF. On admission he presented with expressive dysphasia, right homonymous hemianopia and worsening right sided weakness. An acute CT scan performed within 24 hours showed established infarct in the left occipital and inferior temporal lobes. Further small areas of established infarct were seen in both cerebellar hemispheres. A small lacune was seen in the right subinsular area. MRA was normal.

He was tested a year and a half after the second stroke. At this time he had a right homonymous hemianopia, right hand and facial weakness and expressive and receptive dysphasia. He reports that the control of his hand and his language ability improved only in the last few months. His speech is fluent and fast but is mostly jargon. Repetition and reading aloud are jargon as well. He seems to have difficulty controlling his speech, and therefore unable to refrain from speaking when asked to do so. This pattern is seen in writing as well: his writing output is jargon and long. IF has severe auditory and reading comprehension difficulties. On a verbal comprehension task, he made only semantic errors, while in the written task he made also phonological errors. Due to his reading disabilities, written inner speech tasks were not performed. It was also impossible to establish correct naming of pictures, hence the pictures rhyming task could not be performed as well. Overall, due to his severe comprehension difficulties and the limited comprehensive output, it was impossible to establish whether IF has inner speech.

In the past IF found MRI scans to be extremely unpleasant and he therefore did not want to have another MRI scan.

DH

DH suffered two left MCA strokes, the first one at the age of 74, and the second one at the age of 77. The cause was suspected to be cardio-embolic in both cases. Prior to his retirement he worked as a driving instructor and examiner. It is important to note that DH is left-handed. The first stroke involved the left posterior temporal region and caused mixed aphasia with some comprehension problems, non-fluent output, word finding difficulty and phonological problems. He also had major problems telling the time and doing anything which has to do with numbers. DH received speech and language therapy and him and his wife report that some of his language difficulties resolved with time. However, formal assessment from that time is not available. When admitted to hospital with the second stroke he presented with right sided arm and facial weakness. An acute CT perfusion scan performed within 24 hours showed reduced blood flow in the left temporal lobe, the insula and the posterior frontal lobe.



He was tested 20 months after the second stroke. At this point the right sided weakness was completely resolved but again he was left with mild aphasia. Both his auditory and reading comprehension are mildly impaired but highly functional. His repetition and reading of long or low frequency words is impaired and there are no signs of apraxia. Speech production is characterised by word finding difficulty with both semantic and phonological errors, and writing is severely impaired. DH can perform inner speech tasks. His overt and inner speech are similarly preserved.

Due to his comprehension deficit it was difficult to assess his apraxia. When performing the movement wrongly, he seemed to misunderstand the instructions rather than being unable to perform the movement due to apraxia. He managed to copy all movements without difficulty. An MRI scan was performed at the time of cognitive and language assessment, revealing established areas of infarction in the left occipital lobe, posterior and middle section of the temporal lobe, and parietal lobes. There was also a minor signal change around the right occipital horn but little other evidence of small vessel disease.

JHH

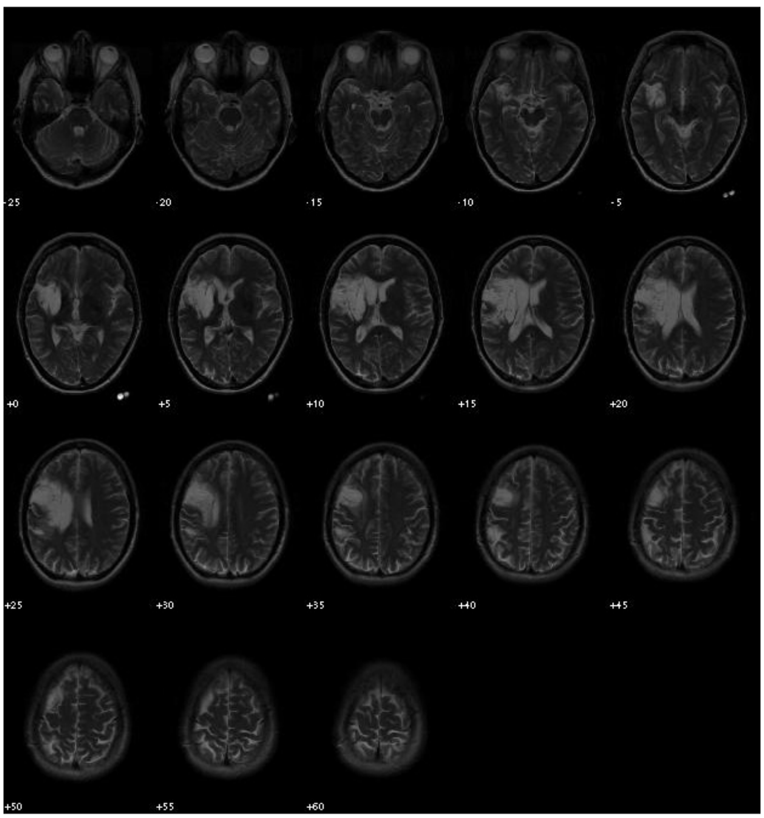
JHH had a left MCA stroke at the age of 62. Prior to having a stroke he worked as a mechanic, having his own business. He initially presented with right hemiparesis and predominantly expressive dysphasia. A CT scan performed within days after the stroke showed a left MCA territory infarct, and low density in paraventricular and subcortical white matter, consistent with small vessel ischaemia. Carotid Doppler was normal. The cause of the stroke is unknown.

When tested 9 years after the stroke, his hemiparesis is completely resolved but he remains with language impairments. Written comprehension, auditory comprehension

and repetition are preserved. Reading aloud is impaired and repetition of non-words is impaired. Spontaneous speech and writing is marked by use of only basic grammatical structures and high level of word finding difficulty.
JHH has a history of Crohn's disease, and have had an abdomen operation in 1991 which used surgical clips thought to be a contraindication for an MRI scan.

GH

GH had a stroke at the age of 41, following a bilateral spontaneous carotid dissection. On admission she presented with dysphasia, dysphagia, and right sided hemiparesis. Her speech was non fluent and she had word finding difficulty, phonological and semantic errors and lack of nouns in spontaneous speech. An MRI scan performed 48 hours post-stroke showed an infarct in the left basal ganglia, insular cortex and adjacent frontal and temporal cortices.



MRA showed left Internal Carotid Artery (ICA) occlusion.
She received Intensive neurorehabiltiaiton.
A year after the stroke she developed stroke related seizures.
When tested 13 months after the stroke her right hand is not functional but she is walking mostly unaided. The right facial weakness is mostly resolved. Her speech production is marked by word finding difficulty and semantic errors. However, she corrects many of her errors. She is helping herself by cueing in writing: when looking for a word she will write the first few letters of that word and then pronounce it out loud. Alternatively, she can also recite the alphabet until reaching the first letter of the desired word. Her speech comprehension is mainly preserved, but she has difficulty

with complex grammatical structures. She is also challenged by semantic distractors: in formal testing she shows hesitations when semantic distractors are present, but if given enough time she does not make semantic errors in comprehension tasks. Her reading is marked by errors, especially for low frequency/low imageability words and she is unable to read non-words. She cannot write complex words or non-words. Sentence writing is impaired as well: she shows severe grammatical errors, especially in confusing prepositions and articles. GH speaks while writing. Remarkably, when doing so she does not make errors when speaking, but her writing does not entirely correspond to her speaking.

Her short-term visual memory is severely impaired but she has no other serious cognitive difficulties. An MRI scan was performed 4.5 months after completion of the cognitive and language assessment. Similar to the acute scan, it revealed a lesion in the left insular region and adjacent frontal and temporal cortices, extending into the inferior frontal gyrus and pre-central gyrus.

TH

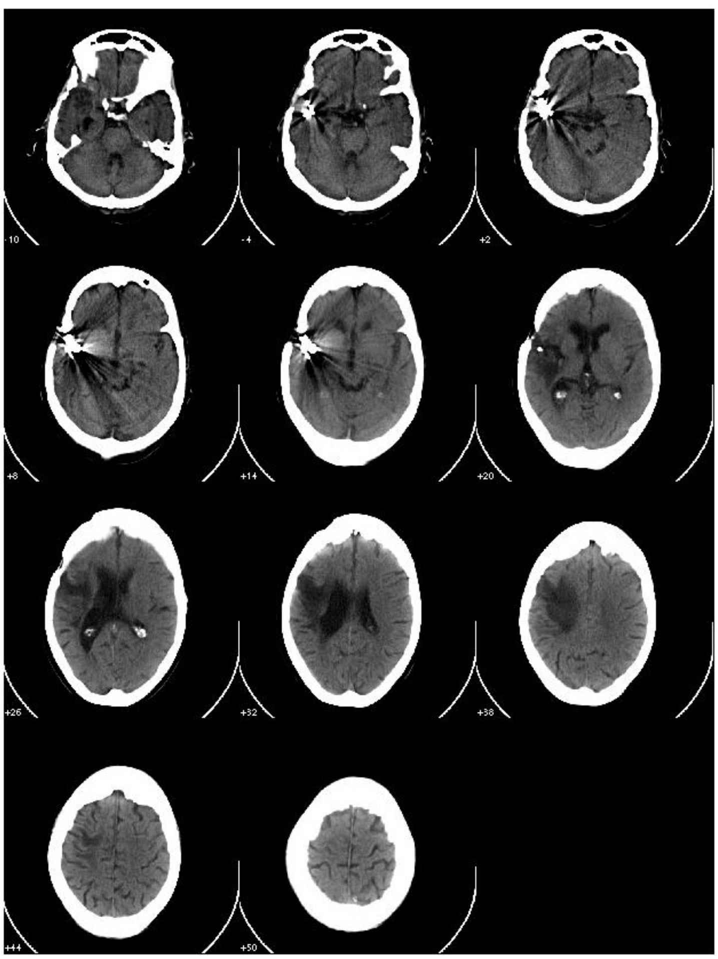
TH is a left-handed woman, who suffered a subarachnoid haemorrhage (SAH) in the left MCA territory at the age of 70, caused by a large aneurysm. Following her stroke she underwent clipping of the aneurysm which caused an ischaemic stroke. She later developed frequent stroke related seizures. Up to her retirement she worked as a Geography teacher. Initial presentation included right upper limb hemiparesis, and speech and swallowing difficulties. A CT scan performed one week after the stroke showed a large artefact from the aneurysm clip which makes it difficult to define the borders of the lesion. An area of moderate low density involving the cortex and underlying white matter of the left lateral frontal lobe was noted. One year later she felt weak and had headaches. She was again admitted to hospital but a CT scan showed no new haemorrhage and ventricle size was reported to be in good size.

She was tested five and a half years after the stroke. Her speech is fluent, slightly agrammatic with hesitations. She has occasional word finding difficulty and makes semantic errors. Same pattern is seen in writing. Her sentence comprehension shows very mild impairment. Repetition and reading of low frequency and non-words is severely impaired.

TH has a mild facial apraxia and weakness. She shows a clear searching behaviour when asked to perform motor commands, and is often hesitant and slow to respond. Her motor output is not accurate; the movement is not clear but recognisable.

She reports that her inner speech is preserved and less effortful than overt speech. On formal testing her inner speech and overt speech of words is similar. However, performing the task with non-words was more revealing. While TH made many mistakes in reading aloud non-words, she performed the homophone task almost perfectly.

TH has lost any numerical ability, including counting for more than 15, telling the time or date and doing simple arithmetic. She also has retrograde and anterograde episodic amnesia. Her memory is preserved only up to her college years. She claims to have forgotten all of her geography knowledge, which she acquired in university. She recognises names of places but cannot tell where they are, point at a map, or give information about the place (for example; population, climate, etc'). She does not recognise places on the map of Britain, except for Kent (where she grew up) and Cornwall. She also testifies that she cannot speak any of the German and French she learned in university.

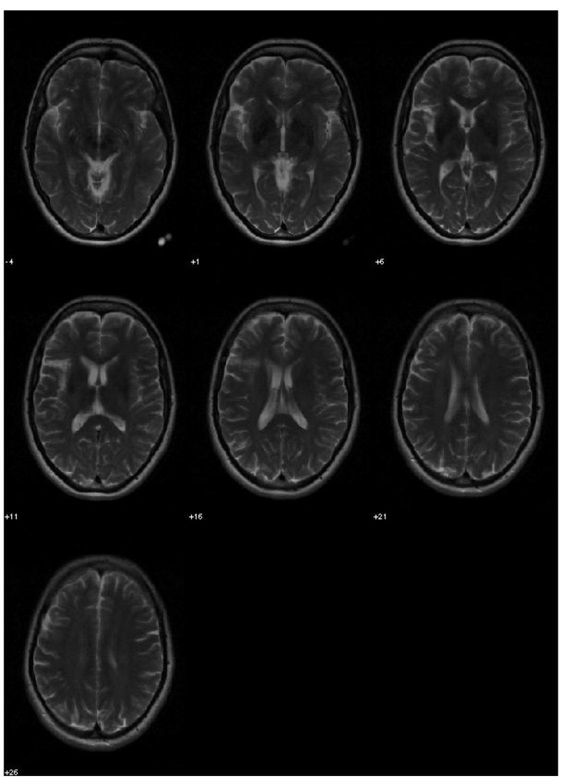


Due to aneurysm clip she could not have an MRI scan and therefore had a CT scan, around the time of the language and cognitive assessment. This showed the clip and related ischaemic area in the left MCA distribution with dilatation of the adjacent ventricle.

LM

LM works full time as a special needs teacher. She had a left MCA ischaemic stroke with an unknown cause, at the age of 48. On admission she was mute but started gaining her voice back within two weeks. Initially she had both semantic and phonological impairments, and could not read, write or repeat non-words. Writing was relatively preserved although she had difficulty with retrieving low frequency, abstract or long words, and at times omitted grammatical words. Errors in writing were rare since her level of pre-production monitoring was high. She showed difficulty with comprehension of complex sentences and impaired verbal short term memory. An acute CT scan performed within 24 hours showed a reduction in cortical attenuation of the left posterior insula as well as prolonged time to peak in the opercular region of the lateral left frontal lobe where there is also reduced cerebral blood flow and minor reduction in cerebral blood volume. Carotid Doppler and MRA were normal.

Right after the stroke she used to be very sensitive to loud noises, which sounded enhanced. Hoovering or door shutting was unbearable. This phenomenon significantly reduced with time but did not diminish. Her voice and accent have changed since her stroke, and this is especially noticeable when she is around new people or in new situations.



When tested a year and 8 months later, comprehension errors still occur. She has difficulty especially with distinguishing active and passive sentences, and with complex sentences. Reading comprehension is better preserved than verbal comprehension. She often shows hesitations between semantic distractors, and occasionally makes semantic errors but no phonological errors. She also has a word finding difficulty. On inner and overt speech tasks using words she performs to a similar level. She reports that the inner speech task is harder. This, however, is not evident from her response time or error rate. However, while she is able to perform the inner speech task with words, she is completely unable to do so with non-words. For these, she has to read them aloud. She reports that she can clearly pronounce words in her head but cannot pronounce non-words in her head at all.

An MRI scan performed 22 months post stroke revealed an infarct in the left inferior frontal gyrus and adjacent insula.

MM

MM had a left ICA ischaemic stroke at the age of 68. A carotid Doppler showed 50% stenosis of the left ICA. Prior to his retirement he worked in the political arena. He used to be right handed but at a late age, after having an injury to his right hand, he discovered he is ambidextrous (similar to his identical twin). However, he has been using his right hand for writing and other functions throughout his life. On admission to the hospital he had right hemiparesis, dysarthria and some problems with speech. Acute imaging is not available since he was admitted to a different hospital. A CT scan was performed 5 months after his stroke, showing a left insular and sub-insular infarct and some small vessel disease. Another scan, performed about two years later, showed substantial global atrophy. MM received no speech therapy after his stroke.

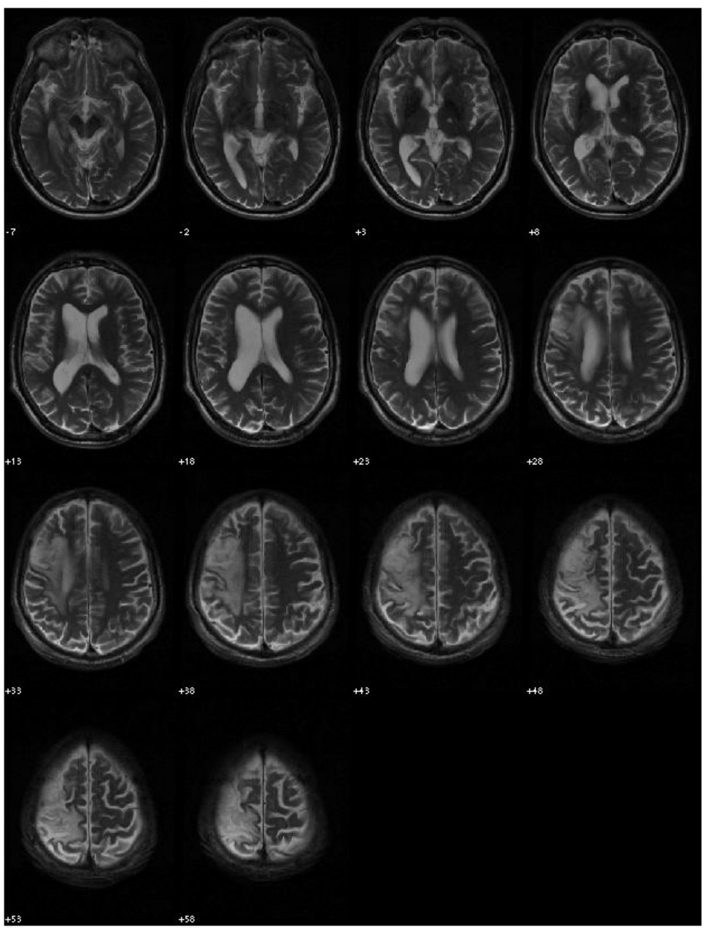
On formal testing, two and a half years after the stroke, MM is mobile, independent and has a very mild aphasia. While his comprehension and repetition are intact, his speech production is slightly impaired, showing word finding difficulty and occasional semantic errors. This is more apparent in spontaneous speech than in formal testing. His writing is impaired; he was able to write only one sentence in 3 minutes, which had both grammatical and spelling mistakes. MM has some difficulties with tasks requiring executive functions, attention and concentration. MM can easily perform inner speech tasks.

MM was recently diagnosed with vascular dementia, and was therefore excluded from the imaging part of the research.

AP

AP is a left-handed man who worked as a painter up to his retirement. He had two ischaemic strokes in the past. The first stroke occurred at the age of 71; causing right hand weakness, right facial weakness, and dysarthria. An acute CT scan performed within 24 hours showed no clear infarct and all symptoms resolved within 2-3 weeks. The second stroke occurred a year later and was caused by stenosis. Initial presentation included a paretic right hand, slight weakness in his right leg and mild expressive dysphasia. An acute CT scan performed within 24 hours and an MRI scan performed two weeks later showed an infarct in the superior part of the left middle frontal gyrus extending as far as the central sulcus. A tiny focus in the right thalamus was also found. MRA showed significant left ICA stenosis.

AP was assessed a year and a half after the second stroke. His comprehension is largely intact, and so is his reading aloud and repetition. He makes rare semantic errors in comprehension tasks and has difficulties with complex sentences. His speech is characterised by occasional word finding difficulty and semantic errors, to which he is not aware. He reports that his main difficulty is in writing, as indeed was found on formal assessment: he can only write very short and simple sentences, and finds it difficult to spell many words.



AP can easily perform inner speech tasks, but his error rate is slightly higher in inner speech than in overt speech tasks.

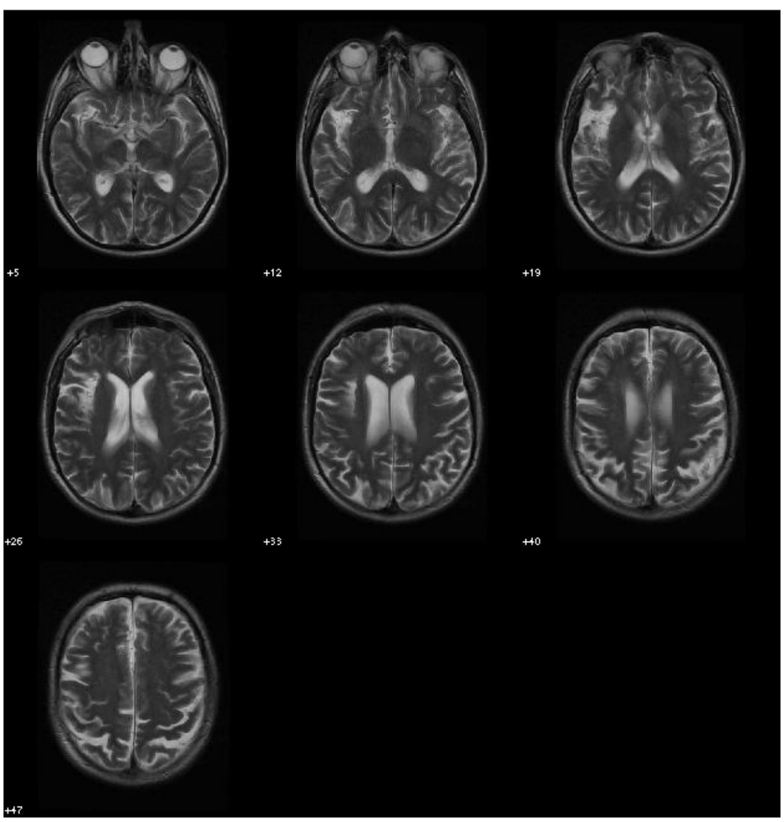
An MRI scan performed at the time of language and cognitive assessment revealed a well established area of signal change and volume loss in the superior left cerebral hemisphere, extending from the left middle frontal gyrus, past the central sulcus, into the post-central gyrus and superior parietal lobe.

CR

CR had a left MCA ischaemic stroke at the age of 59, caused by a PFO and septum aneurysm. He works as a postman and is also self-employed. Initially he was unable to speak, and within a few days he started gaining his speech back but this was non-fluent and anomic. He had no other neurological symptoms and was seen by a speech and language therapist. An acute MRI scan performed within 24 hours showed a left inferior parietal infarct. MRA was normal. He was tested 14 months after the stroke, when his speech is at ceiling on formal testing. Spontaneous speech shows very mild word finding difficulty and some rare semantic errors. He reports that he had dyslexia prior to his stroke which influenced his reading abilities. However, this was never formally diagnosed. He declined an MRI scan.

DR

DR had a left MCA ischaemic stroke at the age of 73, a cause for it was not established. He initially presented with mild expressive aphasia and right facial weakness which fluctuated in severity and resolved within 24 hours. He had no sensory or cognitive deficits and was not seen by a speech and language therapist. An acute CT scan performed within 24 hours showed no clear infarct. Carotid Doppler and MRA were both normal. He was formally tested 11 months after the stroke. On assessment he

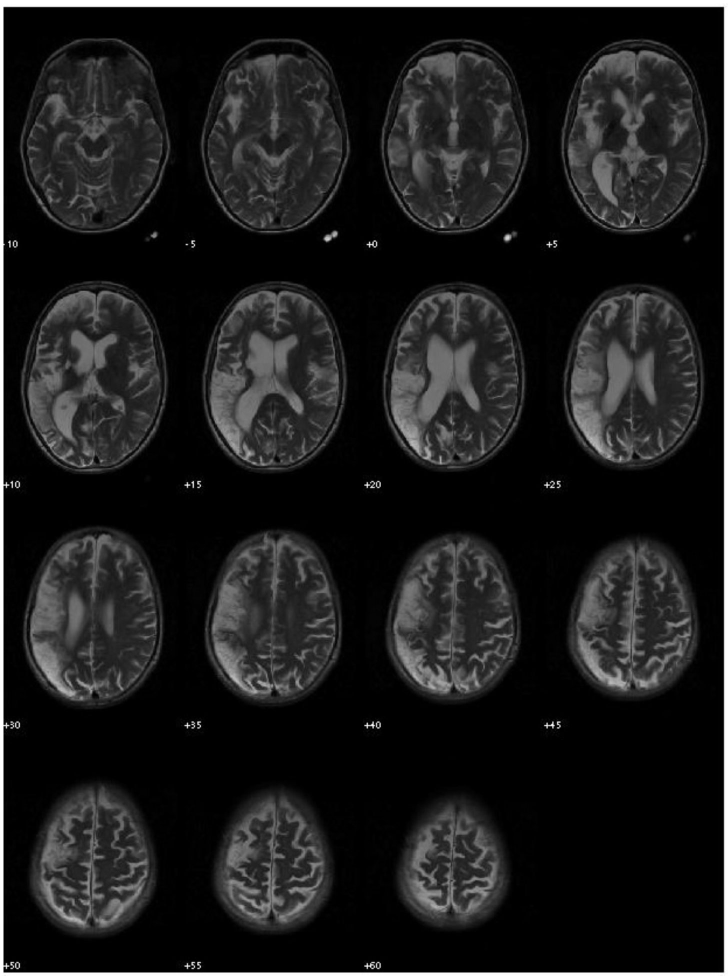


performs at ceiling on almost all tasks. He has difficulty with fluency tasks (producing 8 words in the semantic fluency task and only 4 in the phonological fluency task). He also shows a very mild impairment in comprehension of written, but not spoken, language. Repetition, reading and writing are preserved. Spontaneous speech has occasional semantic errors.

He had a coronary stent inserted in 1999 and therefore he could not be scanned at 3T MRI. Instead he was scanned in the hospital's 1.5T MRI scanner. Unfortunately, a T1 sequence was not performed and so he could not be included in the VLSM/VBM studies. This scan, performed 2 months after testing, showed a well defined and longstanding infarct in the left MCA territory, including the insular cortex and adjacent post-central gyrus.

WS

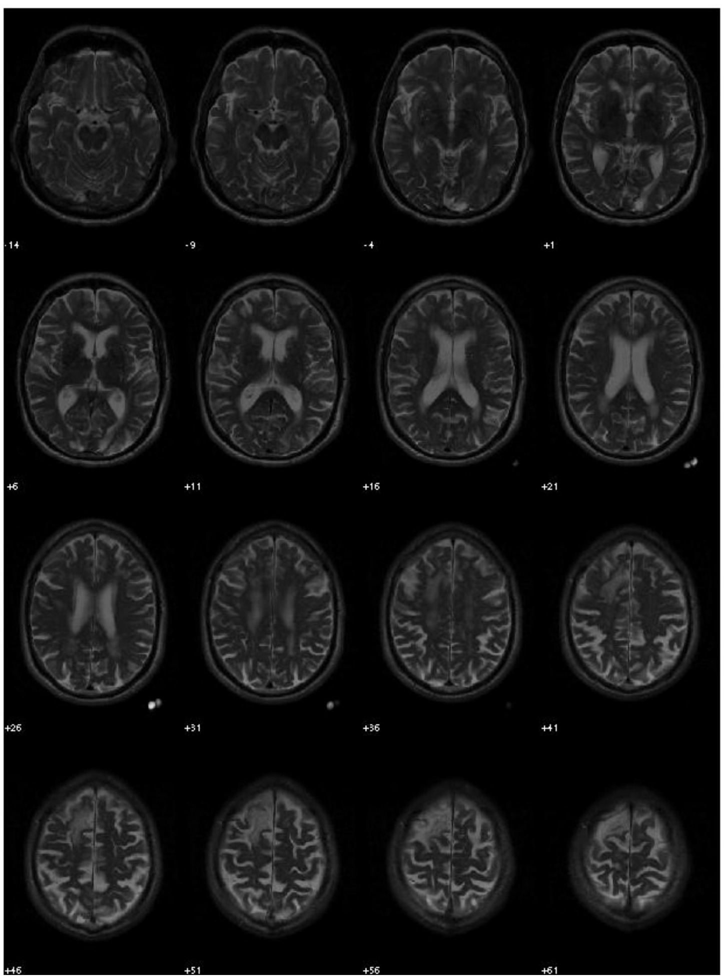
WS had a left MCA ischaemic stroke at the age of 49, two days after a knee replacement surgery. Lately she developed stroke related seizures which are controlled with medication. She has a background history of Systemic Lupus Erythematosus (SLE), osteoarthritis and osteoporosis. Initially she presented with reduced speech and movement in the right arm and leg. She was able to follow commands but not consistently. CT scans were performed within 24 hours and 4 days post-stroke. The scans showed an infarct in the left insular cortex. Carotid Doppler was normal.



She was tested 2 years after the stroke. Her speech output is severely impaired. Comprehension of written and spoken words is preserved. However, comprehension of complex sentences in both modalities is impaired. Reading aloud, repetition and naming all show frequency and word length effect. Although she is able to repeat non words, she is unable to read them aloud. Naming pictures is better performed than predicted based on her spontaneous speech. While spontaneous speech is comprised of mainly yes, no and facial expressions, naming was relatively good, with 17 out of 24 pictures named correctly. Errors comprised of mainly phonological errors. An MRI scan performed at the time of language and cognitive testing showed a lesion in the left insula, middle and superior temporal gyrus, extending superiorly into the post- and pre-central gyri, and posteriorly into the parietal and occipital lobes.

JS

JS is a retired engineer and pilot, and mostly left-handed. He had two ischaemic strokes: the first stroke occurred at the age of 68, and affected the right medial occipital hemisphere. Accordingly, the main symptom was alteration of the visual field, which was resolved shortly after the stroke. Ten years later, at the age of 78, he had a second stroke, affecting the left anterior frontal region, due to AF. This resulted in mild expressive aphasia which was partially resolved within 24 hours. JS also reports that he used to have a phenomenal visual memory, which has severely deteriorated after his second stroke. However, he suspects that some deterioration



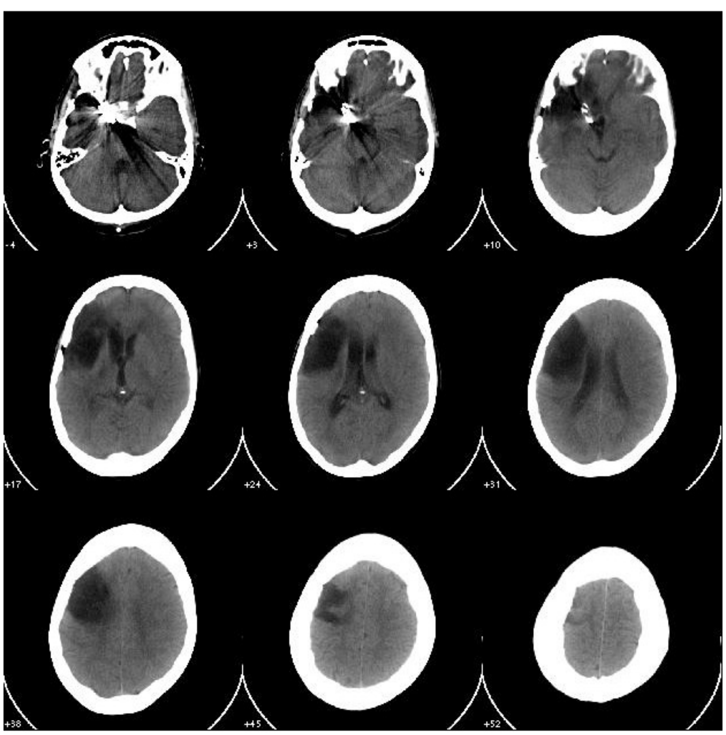
might have started prior to the second stroke. An acute CT scan performed within 24 hours showed a background of small vessel disease and a left anterior frontal infarct, as well as the well established right medial occipital infarct. Carotid Doppler was normal.

He was tested 8 months post-stroke. His speech is almost completely recovered. Every now and then he makes semantic errors (‘sphinx’ for ‘pyramid’, ‘Austria’ for ‘Italy’). Fluency is also affected and sometimes he finishes sentences with vague references such as ‘this and that’. He reports that his language is mostly affected when he is in a hurry, under pressure or not focused enough. He also says that he encounters problems that are semantic in nature, such as picking a spoon instead of a knife or opening wrong drawers. On formal testing he performs almost at ceiling on all tests, except for the fluency tasks (5 words in the phonological fluency task and 9 in the semantic fluency task). His visuo-spatial abilities are somewhat poor but other cognitive functions are preserved.

An MRI scan performed one month after the cognitive and language assessment, showed the same results as the previous CT scan, namely, a background of small vessel disease, and infarcts in the area of the left anterior middle frontal gyrus, white matter of the frontal lobe and right medial occipital region.

AS

AS had an ischaemic stroke at the age of 47 following an elective clipping of a large



ICA aneurysm. She works for a nature conservation organisation.

On admission to the hospital she had a right homonymous hemianopia, right hemiparesis and aphasia. She was discharged from the hospital after 4 months, and at the time had intact comprehension, word finding difficulty, and was able to walk without aid. The hemianopia completely resolved within a few months.

She was tested 10 months after the stroke. Her spontaneous speech is non-fluent, with severe word finding difficulty and lacks appropriate prosody. On naming she makes both phonological (‘strew’ for ‘screw’) and semantic (‘elephant’ for ‘whale’, ‘earth’ for ‘moon’) errors. Her auditory and reading comprehension is intact. In repetition and reading aloud she makes errors only in non-words. A rare error, of the type seen in deep dyslexia, was seen in the reading task (reading ‘horse’ as ‘cat’). On formal assessment her performance is almost at ceiling level. Her difficulties can be seen mainly in the oral and written picture description tasks.

Her inner speech for non-words is mostly intact and is better preserved than overt speech. While often AS is aware of her reading errors, she can rarely correct them, and will give up after a few attempts. Most of her reading errors appear at the end of the word, and are of the type of lexicalisation (reading non-word as a word) or simplification. For words, her reading aloud and inner speech are intact and her performance on the overt tasks was flawless. She finds it difficult to do inner speech tasks and often cannot refrain from verbalising aloud. Lastly, phonological discrimination is impaired for specific close sounds, such as ‘k’, ‘t’ and ‘g’.

Due to the aneurysm clip she could not have an MRI scan and therefore had a CT scan. This scan, performed 2 months before the completion of the language and cognitive assessments, showed an artefact from the aneurysm clip and a mature left MCA territory infarct.

PT

PT had a left MCA ischaemic stroke at the age of 63 due to AF. Up until his retirement he worked as a physical education teacher and swimming pool manager. His initial presentation was of severe global aphasia, right homonymous hemianopia and right hemiparesis. An acute CT perfusion scan performed within 24 hours showed a reduced cerebral blood flow and cerebral blood volume, as well as prolonged time to peak, in the posterior left temporo-parietal region. Carotid Doppler was normal. He initially underwent speech therapy, which had to be stopped due to personal reasons.

He was tested 23 months post-stroke. He is completely recovered from the hemiplegia. His main problem is of speech production. He has severe word finding difficulty and repetition is impaired, mainly for low frequency words. His word comprehension is intact but he has comprehension problems involving certain grammatical structures and relation words. In spontaneous speech he shows mainly phonological errors, and in comprehension, mainly semantic errors. He cannot write at all, but can read well, as long as the font is large and format is 'aphasia friendly'. In some cases he is able to spell aloud words which he cannot write. He can recognise letters by their names and sounds, and can write individual letters by their name and sound as well. He can also copy letters into capitals and into small print. Lastly, he is also capable of copying whole words. However, he is completely unable to write whole words.

He is aided by recitation for speech production. For example, to say a number he will count from 1 until reaching the required number. The same strategy is used for naming days, months, etc. He performed better on the homophone and rhyming task when reading the words aloud, than when using inner speech, but the latter was above chance as well. When using non-words, his performance on the inner speech homophone task was at chance level. When reading aloud, his answers matched his reading, but the performance was poor, reading 75% of the non-words wrong, by producing lexicalisation or other errors. PT declined an MRI.

NET

NET is a retired postman who had a left MCA ischaemic stroke at the age of 63, due to AF. On admission to the hospital he presented with right sided hemiparesis, dysphasia, dysphagia and a visual inattention to the right. A CT scan showed an extensive left MCA territory infarction. In the first few days after admission his mobility improved significantly. Swallowing returned to normal after 10 days. On discharge communication difficulties were still severe and affected by both apraxia and aphasia. Comprehension improved and he could follow two-stage commands.

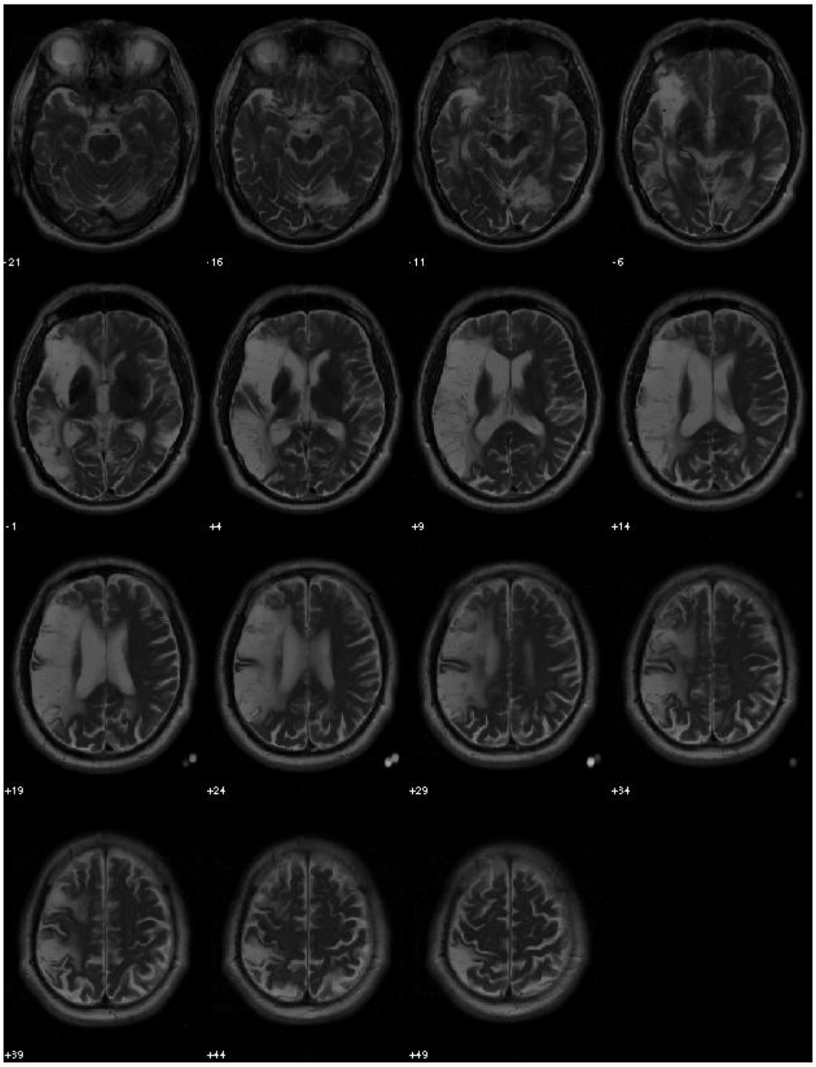
He was tested 7 years and 3 months post-stroke. Semantic memory and recognition are preserved. While both written and auditory comprehensions are just above the cut-off score for impairment level, it is clear that he understands conversations well. He performed well on the auditory paragraph comprehension, showing that when given sufficient context, comprehension is satisfactory.

His main deficit is that of speech production. He can cue himself both semantically (by writing associations) and phonologically (by looking for the right letter on a letter chart, or writing it down), and makes both semantic and phonological errors in speech and writing. He often has access to the initial letter of the word he cannot say or write. However, when unable to retrieve the initial syllable or the entire word, a long cue is needed (one syllable is not enough as a phonological cue). He can often write words which he cannot say. He makes perseverations in writing. He has numerous spelling mistakes mainly involving consonants. However, he is aware of his mistakes and will usually indicate them by saying: ‘not quite’. Interestingly, he adds ‘s’ at the end of words both when speaking and when writing.

When writing to dictation he had difficulty in writing the word, but when shown the object (e.g. pen, man) he can write the word. After noticing this phenomenon, he could immediately use it: when asking to write the word ‘yacht’ he testifies that he imagined a yacht, and wrote it correctly. Lastly, his repetition is severely affected by his apraxia.

He could not perform the inner speech tasks. His phonological discrimination is flawless and when words are said aloud, he can correctly determine if they rhyme or whether they are homophones, or not. However, without hearing them aloud he cannot make the judgement. When doing a rhyming task with pictures he would sometimes write down the word and then try to make a judgement, but he understands that this strategy is not informative. His verbal memory is impaired. Even when given the word corresponding to a presented picture, he is unable to retrieve it a few seconds later.

An MRI scan performed at the time of language and cognitive assessment showed an extensive established infarct in the left MCA territory, involving the middle and superior, frontal and temporal, lobes, and the parietal lobe, as well as a smaller infarct in the right occipital lobe.


-21
-16
-11
-6
-1
+4
+9
+14
+19
+24
+29
+34
+39
+44
+49

Appendix 5: Behavioural Tests

1. Comprehensive Aphasia Test (CAT)

The CAT is a comprehensive, standardised and validated test, which is made to be used with British English speakers. It includes cognitive tests and covers all modalities of language. Scores obtained can be compared to the standardised scores.

1.a. The Cognitive Battery

1.a.a. Line bisection

To evaluate neglect, subjects are asked to estimate the centre of each of 3 lines and to mark it with a pen.

1.a.b. Semantic memory

Subjects are asked to match pictures based on their semantic properties. The target picture appears in the middle of the page and four options appear around it. The subject is asked to choose the picture which is most closely related to the target. the task contains 10 trials.

1.a.c. Word fluency

Subjects are asked to think of as many words as possible, in one minute. In the semantic fluency task, the subject needs to list words which belong to the category ‘clothes’. In the phonological fluency task, the subject is asked to come up with words starting with the letter ‘S’.

1.a.d. Recognition memory

In each of 10 trials, subjects are asked to indicate which of four pictures appeared before, in the semantic memory task. This task is administered about 3 minutes after the semantic memory task.

1.a.e. Gesture object use

Subjects are presented with 6 pictures of objects and are asked to show with their hands how they might use the object.

1.a.f. Arithmetic

Subjects are asked to solve 6 equations, including subtraction, addition, multiplication and division.

1.b. The Language Battery

1.b.a. Comprehension of spoken words

Subjects are given a spoken word and are asked to point to the corresponding picture. They have a choice of 4 pictures, which include the target, a phonological distractor, a semantic distractor and a non related distractor. This task has 15 trials.

1.b.b. Comprehension of written words

This task is identical to the previous one, only this time the target is a word written in the middle of the page.

1.b.c. Comprehension of spoken sentences

Subjects are read out a sentence and are asked to point to one of four pictures which best goes with the sentence. The task has 16 trials.

1.b.d. Comprehension of written sentences

This task is identical to the previous one, only this time the sentences are written in the middle of the page, instead of spoken out loud by the examiner.

1.b.e. Comprehension of spoken paragraphs

The subjects are read two short stories. At the end of each story they are asked 4 yes/no questions. If unable to speak the subject can point at the written words to indicate his or her answer.

1.b.f. Repetition of words and non-words

Subjects are asked to repeat words which are read by the examiner. This include 16 short words, 3 complex words (with prefixes and affixes), and 5 non-words.

1.b.g. Repetition of sentences

Subjects are asked to repeat sentences with increasing length. The shortest sentence has 3 content words and the longest sentence has 6 content words. Phonemic errors, apraxic errors and dysarthic distortions are accepted as correct because the aim of the test is to assess sentence span only.

1.b.h. Naming objects

Subjects are asked to name 24 pictures of nouns.

1.b.i. Naming actions

Subjects are asked to name the action depicted in each of 6 pictures. They are prompted with the question: ‘What is he doing? He is…’ or ‘What is she doing? She is…’

1.b.j. Spoken picture description

Subjects are asked to describe a complex picture. If areas of the pictures are omitted from the description the examiner prompts the subject to describe it by asking: 'What about this?' The verbal description is recorded and later transcribed and coded. The description is limited to 3 minutes.

1.b.k. Reading words and non-words

Subjects are asked to read aloud a list of words and non-words. The list includes 24 simple words, 3 complex words, 3 function words, and 5 non-words.

1.b.l. Writing: copying

Subjects are asked to copy letters. 5 letters are copied from upper case to upper case and 6 different letters are converted from lower case to upper case. The subjects are then asked to copy 3 words, converting the letters from lower case to upper case.

1.b.m. Writing picture names

Subjects are asked to write the names of 5 pictures of nouns.

1.b.n. Writing to dictation

Subjects are asked to write 5 words to dictation. This includes a short concrete word ('man'), an irregular concrete word ('yacht'), an abstract word ('idea'), a morphologically complex word ('undrinkable'), and a non-word ('blosh').

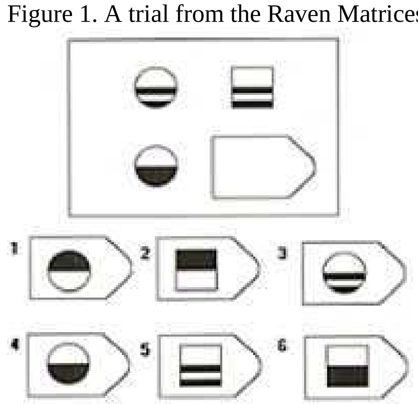
1.b.o. Writing picture description

Lastly, subjects are presented again with the same picture given for the verbal description and are asked to describe the picture in writing. Time allowed for this task is also 3 minutes.

2. Raven Matrices

Subjects are presented with a pattern with a missing piece. They are asked to indicate which one of 6 or 8 options presented below will complete the pattern. Subjects completed blocks A, B and D, each containing 12 trials. For an example see figure 1.

Figure 1. A trial from the Raven Matrices.



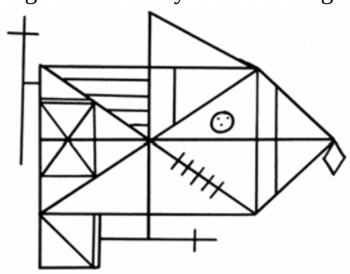
3. Apraxia Battery for Adults

Subjects performed sub-tests 1 and 3 from the Apraxia Battery for Adults. In sub-test 1, they are asked to repeat combination of syllables. In sub-test 3, limb and oral apraxia are examined by asking subjects to perform various hand and oral motor actions.

4. Rey–Osterrieth Figure

Subjects are presented with the Rey–Osterrieth figure (below) and are asked to copy it, without any time limit. After half an hour in which they perform other unrelated tests, they are asked to draw the figure from their memory.

Figure 2. The Rey–Osterrieth figure



5. Brixton Test

The Brixton Spatial Anticipation Test examines executive functions. Subjects are asked to recognise the pattern of movement of a single dot and based on its previous moves, to guess where it will go next. The pattern is changed every few pages and subjects are required to notice this, and change their guesses accordingly. They are not given any indication as to when the changes in the pattern occur.

6. ACE-R

Subjects performed the sub-tests of the ACE-R which examine visuo-spatial and perceptual abilities. These include copying overlapping pentagons, copying a wire cube, drawing a face of a clock, counting dots without pointing at them and identifying incomplete letters.

Appendix 6: fMRI Study - Stimuli

Word Condition

Rhyming-pairs		
Primary rhymes	Bear	Chair
	Pear	Square
	Bow	Dough
	Brain	Crane
	Heart	Dart
	Door	Drawer
	Fly	Tie
	Fox	Socks
	Leek	Beak
	Bowl	Mole
	Egg	Peg
	Thumb	Drum
	Soap	Pope
	Seal	Wheel
	Shoe	Screw
	Suit	Boot
	Whale	Nail
	Chain	Plane
Rhymes with repeated pictures	Square	Bear
	Chair	Pear
	Chair	Square
	Crane	Train
	Tie	Eye
	Eye	Fly
	Socks	Box
	Snail	Whale

Non-rhyming pairs (derived from primary rhymes which are not repeated)		
Similar ending	Dart	Suit
	Wheel	Bowl
	Boot	Heart
	Pope	Shoe
	Dough	Egg
	Seal	Mole
Similar beginning	Screw	Soap
	Bow	Beak
	Drum	Door
No similarity	Drawer	Peg

Picture Condition

Rhyming-pairs		
Primary rhymes	Chair	Bear
	Square	Pear
	Dough	Bow
	Crane	Brain
	Dart	Heart
	Drawer	Door
	Tie	Fly
	Socks	Fox
	Beak	Leek
	Mole	Bowl
	Peg	Egg
	Drum	Thumb
	Pope	Soap
	Wheel	Seal
	Screw	Shoe
	Boot	Suit
	Nail	Whale
	Plane	Chain
Rhymes with repeated pictures	Bear	Square
	Pear	Chair
	Square	Chair
	Train	Crane
	Eye	Tie
	Fly	Eye
	Box	Socks
	Whale	Snail

Non-rhyming pairs (derived from primary rhymes which are not repeated)		
Similar ending	Boot	Dart
	Bowl	Seal
	Suit	Heart
	Bow	Screw
	Mole	Shoe
Similar beginning	Drawer	Drum
	Peg	Pope
	Door	Dough
No similarity	Soap	Wheel
	Egg	Beak

Practice Block - Words

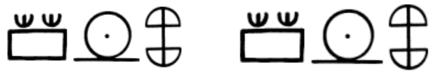
Rhyming pairs	gate	plate
	ant	plant
	cat	hat
	tram	pram
	log	dog
	sun	bun
	bag	flag
	moon	spoon
	snake	cake
	wall	ball
Non-rhyming pairs	wall	tram
	log	bag
	hat	ant
	dog	flag
	plant	cat
	plate	snake
	ball	pram
	sun	moon
	cake	gate
	bun	spoon

Practice Block – Pictures

Rhyming pairs	hat	cat
	ball	wall
	spoon	moon
	flag	bag
	plant	ant
	pram	tram
	bun	sun
	plate	gate
	cake	snake
	dog	log
Non-rhyming pairs	wall	pram
	bag	dog
	gate	snake
	tram	ball
	plate	cake
	log	flag
	cat	ant
	plant	hat
	sun	spoon
	moon	bun

Word Base Line Condition (Examples)

Similar Pair (Yes)



Dissimilar Pair (No)

 ****

Picture Base Line Condition (Example)

Similar Pair (Yes)

Dissimilar Pair (No)

Appendix 7: Abbreviations

A1 – Primary Auditory Cortex
ABA - Apraxia Battery for Adults
ACC – Anterior Cingulate Gyrus
ACE-R - Addenbrooke's Cognitive Examination – Revised
AF - Atrial Fibrillation
AVM - Arteriovenous Malformation
BA – Brodmann Area
BOLD – Blood-Oxygen-Level-Dependence
CAT - Comprehensive Aphasia Test
CIT – Constrain Induced Therapy
CSF – Cerebro-Spinal Fluid
CT - Computed Tomography
CTA - Computed Tomography Angiogram
DLPFC – Dorsolateral Prefrontal Cortex
DTI – Diffusion Tensor Imaging
DVT - Deep Vein Thrombosis
EEG – Electroencephalography
EPI - Echoplanar Images
FA – False Alarm
fMRI – Functional Magnetic Resonance Imaging
FOV - Field of View
GM – Grey Matter
HRF - Haemodynamic Response Function
ICA – Internal Carotid Artery
IFG – Inferior Frontal Gyrus
IPL – Inferior Parietal Lobule
ISI - Inter-Stimulus-Interval
ITG – Inferior Temporal Gyrus
MCA - Middle Cerebral Artery
MEG – Magnetoencephalograph
MFG – Middle Frontal Gyrus
MNI - Montreal Neurological Institute
MRA - Magnetic Resonance Angiogram
MRI – Magnetic Resonance Imaging
MTG - Middle Temporal Gyrus
PALPA - Psycholinguistic Assessments of Language Processing in Aphasia
PET – Positron Emission Tomography
ROI - Region of Interest
RT - Reaction Time/Response Time
SAH - Subarachnoid Haemorrhage
SMA – Supplementary Motor Area
STS – Superior Temporal Sulcus
STG – Superior Temporal Gyrus
TE - Echo Time
TIA – Transient Ischaemic Attack
TOT – Tip of the Tongue
TR - Repetition Time
TTP – Time to Peak
VBM – Voxel Based Morphometry
VLSM – Voxel Based Lesion Symptom Mapping
VWFA - Visual Word Form Area
WM – White Matter