Phonetic Encoding, Verbal Working Memory and The Role of Broca's Area

Marina Papoutsi

Doctor of Philosophy

University of Edinburgh

2008

Abstract

Even though Broca's area has been associated with speech and language processing since the 19th century, the exact role that it plays is still a matter of debate. Recent models on the neuroanatomical substrates of language have assigned Broca's area to different processes: syllabification (Indefrey and Levelt 2004), articulatory code storage (Hickok and Poeppel 2004) and verbal working memory (Chein and Fiez 2001; Chein et al. 2002). The subject of this doctoral dissertation, is to examine language production and disambiguate the role of Broca's area. This issue was addressed in a series of functional magnetic resonance imaging studies (fMRI) involving speech production, where the phonological properties of pseudowords were manipulated in a way that differentiated between syllabification and articulatory code generation. The load on verbal working memory was also changed. The behaviour of Broca's area was then examined in response to these manipulations to determine the dependence of the observed results on the different levels of processing and verbal working memory.

The results from the present studies suggest that the dorsal premotor cortex has a consistent role in articulatory code generation irrespective of verbal working memory demands. In contrast, Broca's area, specifically Brodmann area 44, showed a main effect of phonetic encoding only during delayed response tasks. Interestingly, area BA44 was also found to be functionally segregated between the dorsal and ventral part. The dorsal part was sensitive to articulatory and phonological load, such as stimulus length. The ventral part on the other hand was sensitive to sub-lexical stimulus properties, but only during delayed response trials. These findings suggest that BA44 is not a homogeneous region, but it is divided into a dorsal premotor and a ventral prefrontal part. These results add another dimension of complexity to the study of Broca's area, its functional segregation and its role in language production.

Acknowledgements

For the completion of this thesis I am grateful to many people for their support and guidance. I am most grateful to my supervisors at the University of Edinburgh, Drs. James Bednar and Martin Pickering, for their confidence in me and continuous support and encouragement. I am also grateful to Dr. Barry Horwitz at the NIDCD/NIH, who has been a wonderful mentor throughout this journey and who has generously welcomed me in his lab. Thank you all so much. It has been a great pleasure for me to work with you and learn from you all.

My thanks also extends to the Edinburgh Neuroinformatics Doctoral Training Centre (DTC) and all those part of it, both students and staff, past and present. Particular thanks goes to Dr. Mark van Rossum, Prof. David Willsaw and Mrs. Pat Ferguson. Their continuous help and support have been invaluable. Credit also goes to the EPSRC and MRC who have provided financial support through the DTC and the Bakalas Brothers Foundation for their generous stipend.

I also would like to thank my colleagues and friends both at the University of Edinburgh and the NIH. In particular, from the NIH I would like to thank Drs. Jacco de Zwart and Martijn Jansma, who have generously shared their expertise and time with me; Drs. Jason Smith, Jieun Kim, Fatima Hussain and Feng Rong who have been a great group to work with and who have taught me a lot; Mrs. Alda Ottley for her precious help and the wonderful company during the many hours of scanning; Dr. Jeff Duyn for generously allowing us to take up some of his scanning slots; and Dr. Allen Braun for being ever so welcoming. Many thanks also go to Dr. David McGonigle who has helped a lot with the initial stages of the thesis; Ms. Elvina Gountouna who has been a great friend, colleague and companion throughout this journey, Ms. Lizelle Bischoff who helped me in more than one way during her time at the DTC and last but not least all the informatics computing support team and their invaluable help.

Finally, my deepest gratitude and thanks go to my family, both close and extended. Thank you all so much for your unconditional love, your continuous support and most of all for being there and helping me grow both physically and spiritually. Special thanks also go to Vasso. Thank you for your loving support during these past few months, for doing all the cooking and cleaning and most of all for being the great sister that you are!

Thank you all!

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(Marina Papoutsi)

Table of Contents

Index of Tables

Illustration Index

As you set out for Ithaca,
hope your journey is long
and full of adventures ...

C. Cavafy 1911

Chapter 1: Introduction

The subject of this thesis is the system of phonological and phonetic encoding and in particular the role of the posterior left inferior frontal gyrus, also known as Broca's area. The first chapter of the thesis includes a description of the theoretical and experimental background that led to the conception and implementation of the present experimental work on this system using functional magnetic resonance imaging. Discussed are some of the current theoretical and computational models on language production, as well as the results from a number of studies in the fields of neurology, neuropsychology, psycholinguistics and neurophysiology. In the recent years, much progress has been made in the study of language production and the identification of its neuroanatomical substrates. Still, there are many questions left unanswered and we are still far from having a clear understanding of the mechanisms behind language production and the precise function of one of its key regions, Broca's area. While many studies have looked into the role of this region, the variety of tasks and, even more so, the variability in the definition of the region itself have led to numerous apparent contradictions. This chapter will set the context for the experiments that will be presented in the next chapters.

1.1 Experimental Framework and Basic Definitions

The overall goal of this thesis is to understand human language production. Within this general framework the focus is on the generation of articulatory codes, the identification of the key anatomical areas, their role and their interactions. The left inferior frontal gyrus (LIFG) and in particular Brodmann area 44, has been shown to play a particularly important role in language production, though the precise details of its function are yet to be resolved and there are many different and often contrasting opinions. The work presented as part of this thesis focuses on disambiguating the role of the LIFG in language production and on providing more information about its function. New findings from the field of neuroimaging suggest that a greater functional segregation of the LIFG exists than previously believed (Chein and Fiez 2001; Chein et al. 2002; Friederici 2002; Molnar-Szakacs et al. 2005). The subject of this thesis was precisely the issue of functional segregation within the LIFG and the significance of such a segregation with respect to phonetic encoding and language production in general. A series of functional magnetic resonance imaging (fMRI) experiments on humans was performed in an effort to answer these questions. The results (as presented in chapters 4, 5 and 6) provide evidence for the existence of a dorsal-to-ventral gradient of functional specialization within the posterior LIFG, Brodmann area 44 (BA44). Consistent with its anatomical location between the premotor and the prefrontal cortex, BA44 seems to be related to both prefrontal and premotor processes. However, the present findings also challenge the hypothesis that the LIFG is the key region underlying phonetic encoding and articulatory code generation. This role seems to be more appropriate for the premotor cortex. However, more details about the experimental findings will be presented in the following chapters. This first chapter will include a description of the framework of this project and the definition of some of the terms that will be used throughout the thesis. Previous research will then be related to the work of the thesis, in order to help the reader understand in more depth both the aim and the significance of this work.

A first step is to clarify what is meant by phonetic encoding or phonological processing, because these terms can have different meanings in different research fields. In it's simplest view, phonology refers to the study of speech sounds and their organization. Traditionally, the phonological unit, also referred to as a phoneme, is an abstract category of speech sounds that allows words to be distinguished and acquire meaning. To use a conceptual analogy, there are many different types of triangles (isosceles, equilateral etc.), yet they are all categorized as triangles, if they fulfil certain fundamental requirements (namely being a polygon with exactly three sides). Phonemes can be thought of similarly. There can be many different pronunciations of /a/ at the phonetic level, but as long as the sound meets certain acoustic frequency requirements it will be perceived as the phoneme /a/ and it will not be confused with e.g. the phoneme /e/.

In language production, phonological representation is thought to be one of the intermediate steps as a speaker proceeds from a conceptual representation of the intended utterance to a spoken articulation. As Levelt and his colleagues put it (Levelt et al. 1999), there is a rift between conceptual and syntactic representations and forming an articulatory plan for this representation. Phonological processing is the first step in this process and it is very strongly linked to the generation of the articulatory codes. More details on the different levels of representation will be reported below, when some of the current models on language processing and its neuroanatomical substrates are described. Current models of language production will be presented along with research in support or against these models, information on the experimental questions that will be addressed in this thesis and details on how they will be addressed.

1.2 From Phonological Codes to Articulatory Scores

As previously mentioned, at the level of the word form, successful articulation requires generating an appropriate motor plan. In this process, an abstract, internal representation of a word, what is referred to as the phonological representation, is transformed into an articulatory representation. This transformation is by no means simple. It involves multiple layers of representation, e.g. phonological and phonetic, and the engagement of a wide cortical network surrounding the Sylvian fissure, before the final form of the articulatory code is generated. A number of computational and theoretical models have been presented to explain this process and to understand the underlying cognitive mechanisms. However, to this day, only very few of those models make specific hypotheses about the cortical regions and neuronal processes that might be taking place. For the purposes of this research the focus will only be on models of language production that include specific hypotheses about anatomical regions, such as the models proposed by Indefrey and Levelt (2000; 2004) and Hickok and Poeppel (2000; 2004; 2007). Of particular interest are the hypotheses that these two models make about the role of Broca's area in the process of generating articulatory codes.

The processes that lead to the generation of an articulatory motor plan are a matter of debate amongst researchers, as is the timing and interaction between these processes (for a review see Goldrick and Rapp 2007). However, it is commonly accepted that syllabic, metrical and featural information (though possibly only the non-redundant features) have been specified in a phonological representation, prior to the generation of a motor plan (Levelt 1999). In extended reviews of studies on word production by Indefrey and Levelt (2000; 2004), it was suggested that in the final stages prior to phonetic encoding and the generation of the articulatory representation (articulatory score), the generated phonological code is spelled out into its different phonemic segments. It is then clustered into syllables and assigned a metrical structure, a process described as syllabification. As syllables are created, they are also rapidly

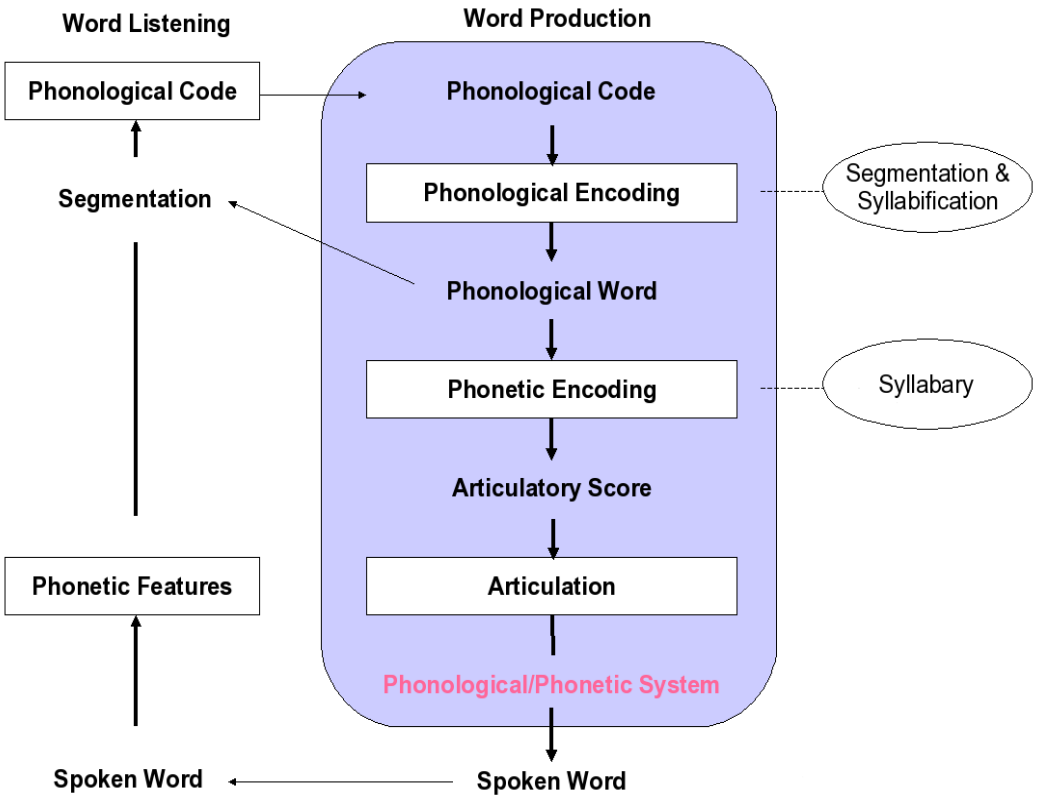

Figure 1: Network of processing components involved in speech production following auditory input. This diagram has been adapted from Indefrey and Levelt (Indefrey and Levelt 2004) to represent repetition of auditorily presented words. Left column: assumed processing steps in word listening. Middle column: core processes of word production. Shown in white boxes are the different processing steps in the system. The arrows describe the direction of the processing, e.g. input or output. For example, a phonological code is the input of phonological encoding. The output is a phonological word. Shown in pink is the name of the system. Right column: examples to clarify the processes described on the middle column. According to the presented model the phonological input code is different from the phonological code generated for output. The arrows connecting the two columns represent feedback.

turned into sequences of motor gestures, also known as gestural or articulatory scores (Browman and Goldstein 1988; see Figure 1 for a diagram of the proposed model). In a pseudoword repetition setting for example, one hears the pseudoword /kɪkˈɛb/ and is asked to reproduce it. After generating the phonological code of the pseudoword and separating it into its phonemes, i.e. /k/ /ɪ/ /k/ /ɛ/ /b/, then the

segments would be clustered into syllables, e.g. [kɪ]-[k'ɛb]. For each syllable the gestural code would then either be retrieved or compiled depending on the syllable frequency of occurrence. Articulation can begin as soon as the first syllable is fully phonetically encoded (Bachoud-Lévi et al. 1998).

In this account of single word production, syllables are the fundamental units in constructing the articulatory representation and it is also assumed that there is a different mechanism in dealing with high and low frequency syllables. Based on the notion that speakers tend to re-use only a small number of syllables and on evidence that pseudowords with high frequency syllables are faster to produce than their low frequency counterparts (Cholin et al. 2006), it has been proposed that the articulatory scores for frequent syllables are pre-compiled and stored in a repository called the "mental syllabary". In contrast, the articulatory representations for less frequent syllables have to be compiled on-line (Levelt and Wheeldon 1994).

Neuroanatomically, the process of generating lexical phonological representations has been associated with the middle and posterior superior temporal gyrus (Ohbayashi et al. 2003; Fiez et al. 1999; Indefrey and Levelt 2000; Hickok and Poeppel 2004), also known as Wernicke's area. In some theories (Zatorre et al. 1996; Poldrack et al. 1999; Burton et al. 2000), it has also been assigned to Broca's area and specifically to the posterior part of the LIFG, roughly corresponding to BA44. This region is thought to be specifically involved in syllabification (Indefrey and Levelt 2000) and sub-lexical processes that require explicit segmentation, such as tasks where subjects perform phonological decisions like phoneme monitoring, phoneme discrimination, or phoneme sequencing (Zatorre et al. 1992; Zatorre et al. 1996; Demonet et al. 1996; Poldrack et al. 1999; Burton et al. 2000). In a proposed model by Indefrey and Levelt (2004), the LIFG is part of a network related to segmenting a retrieved phonological word, while the premotor cortex (Brodmann area 6) is responsible for compiling and storing the motor codes for the individual syllables. Hence, according to this view, the premotor cortex is identified as the location of the mental syllabary rather than the LIFG.

In recent review papers, Hickok and Poeppel (2004; 2007) followed a different approach for understanding linguistic processes. The Hickok and Poeppel model was inspired by the theory of the "mirror neuron system" (MNS) and the idea of sensory-motor integration (di Pellegrino et al. 1992; Rizzolatti and Arbib 1998; Rizzolatti and Craighero 2004). According to the MNS theory and its extension for language (Rizzolatti and Arbib 1998), there is a common interface between speech perception and production, which also facilitates phonemic-to-articulatory code translation and is in agreement with the "motor theory of speech perception" (Liberman and Mattingly 1985). According to the motor theory of speech perception, successful understanding and communicational parity[1] require a form of sensory-motor mapping that will encode the lexical - or any other - sensory input to the listener's own motor system (Liberman and Mattingly 1985). This theory is in agreement with the MNS theory and its extension to include language, while evidence from research on the mirror neuron system also provide an anatomical substrate for the sensory-motor mapping. Broca's area is considered to be part of the sensory-motor integration interface and directly involved in the generation or retrieval of the articulatory codes. Following a computational model of speech production, the proposed role of the posterior part of Broca's area (along with the ventral premotor cortex) is to hold a "speech sound map", i.e. representations of phonemes or frequent syllables and their associated motor programs (Guenther et al. 2006).

The concept of the speech sound map is similar to the idea of the mental syllabary presented above, in the model proposed by Indefrey and Levelt (2000). Where the two theories differ is in the role of the posterior part of Broca's area. According to Hickok and Poeppel, the role of Broca's area is phonetic encoding and the generation of the articulatory scores, since it serves as a store for articulatory representations. In contrast, Indefrey and Levelt argue that the role of Broca's area is to support syllabification and phonological encoding, which are processes that are a step before the generation of the articulatory codes.

1) Communicational parity is the situation where the speaker and the listener share a common knowledge.

To support their claims, Indefrey and Levelt referred to evidence that activations in Broca's area are independent of whether the task requires overt or covert response and therefore not directly related to the generation of articulatory codes. Based on their model, segmental processing and syllabification are the last common steps in the process of word production and prior to generating the articulatory code. However, as the authors themselves have acknowledged, it is still possible that in cases of covert response the articulatory code is retrieved. Whether the articulatory code will be retrieved or not during covert speech tasks seems to be highly dependent on the task instructions and not the task response demands. For example, when covert repetition is defined as covert rehearsal of the target stimulus or when the "phonological loop" is activated (Baddeley 2003), then it is assumed that the complete articulatory code is generated (Indefrey and Levelt 2000). On a further note, based on the theory of sensory-motor integration during speech (Hickok and Poeppel 2000), as well as the motor theory of speech perception (Liberman and Mattingly 1985), articulatory codes could be retrieved/compiled not only during word production, but also during perception. This effect is particularly highlighted by studies using transcranial magnetic stimulation (TMS) of the motor cortex and positron emission tomography (PET) which showed that speech-related motor muscles and cortical regions have increased excitability during speech perception (Fadiga et al. 2002; Pulvermüller et al. 2006). Therefore, it is possible that the articulatory code is generated independently of the specific task demands on overt response[2].

From what has been reviewed so far, views on the function of Broca's area cover a very wide range. To anyone who has studied this region, the contrasting views come as no surprise, since this area appears not only functionally, but also anatomically complex. The following sections provide a description of the main anatomical and functional characteristics of the region in an attempt to gain a better understanding of

2) For a review on evidence in support of the engagement of the motor system during speech perception see Galantucci et al. (2006).

the previously discussed models of word production and the exact role that Broca's area plays in the process.

1.3 Broca's Area: Anatomy and Function

In 1861 the French surgeon Pierre Paul Broca made the first presentation of the case of Monsieur Leborgne, a French worker who had lost almost all ability to speak apart from saying the syllable “tan”. Dr. Broca referred to his condition as “aphemie”, currently known as “aphasia” (Broca 1861). Monsieur Leborgne had an extensive lesion in his left hemisphere, which included, but was not limited to, the posterior LIFG. According to Dr. Broca and based on observations from other patients, this region was the “seat of speech” and damage to this area would result in severe motor aphasia. Because of the work and discoveries that Dr. Broca made on the study of this area, today the posterior part of the LIFG is also referred to as Broca's area, while it has also retained its status as one of the most important nodes in the brain network of language and communication.

However, in the recent years it has become more evident that the area is far from being only specialized for speech. Instead, it has a more general role extending beyond speech to include working memory, sensory motor integration, motor sequencing etc. The exact role of the region is yet unclear and theories have also pointed to both a functional (Chein et al. 2002) and an anatomical (Amunts et al. 1999) segregation within the area. That is, it has been proposed that different parts of the area serve different functions, which can explain the region's seeming multi-functionality. These theories were inspired by both functional and anatomical data.

1.3.1 The Anatomy of Broca's Area

With respect to anatomy, Broca's region is situated in the ventral posterior IFG.

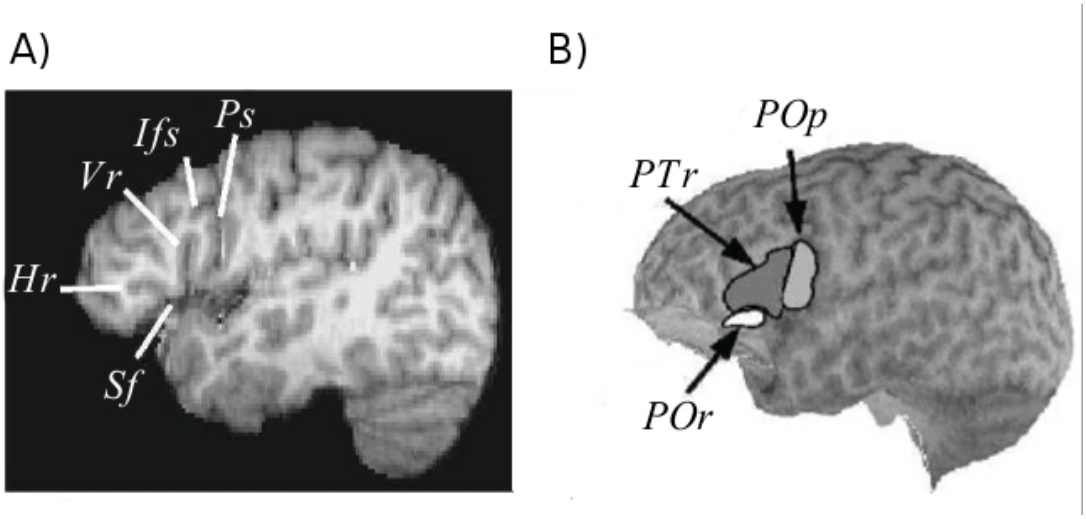

Figure 2: Depiction of the gyral anatomy of the posterior inferior frontal gyrus. Shown in (A) is a sagittal view of the left hemisphere of the human cerebral cortex, where one can see Broca's area. The annotations highlight major anatomical landmarks that define Broca's area and its three parts. The pars opercularis (referred to in (B) as POp) is defined rostrally by the vertical ramus of the Sylvian fissure (Vr), caudally by the inferior segment of the precentral sulcus (Ps), dorsally by the inferior frontal sulcus (Ifs) and ventrally by the Sylvian fissure (Sf). Anterior to the POp is the pars triangularis (referred to in (B) as Ptr), which is defined dorsally by the Ifs and ventrally by the horizontal ramus (Hr) of the Sf. Finally, the pars orbitalis (referred to in (B) as POr) is ventral to the PTr and extends to the lateral orbital sulcus. It is ventrally limited by the Sf. Shown in (B) is another rendering with the three parts highlighted in different shades of gray. This image has been adapted from Devlin et al. 2003.

When most people refer to this area they also associate it with the left hemisphere. There is also a right hemisphere anatomical homologue of the area, although functionally speaking the right hemisphere IFG seems to be involved in different aspects of cognitive processing, such as pitch perception (Zatorre et al. 1992). Anatomically, Broca's area is located between the premotor and prefrontal cortex and it is approximately defined by the precentral sulcus (posterior border), the inferior frontal sulcus (dorsal border) and the anterior horizontal ramus of the Sylvian fissure (inferior border) (see Figure 2 for a depiction; (Devlin et al. 2003).

Despite the fact that the posterior, dorsal and inferior borders of Broca's area are

quite well defined, the anterior borders are more disputed. Therefore, it is unclear how far into the prefrontal cortex Broca's area extends to. By means of its gyral patterns, the posterior IFG is divided into the pars opercularis, the pars triangularis and the pars orbitalis. By means of the cytoarchitectonic laminar patterns, it corresponds to Brodmann areas BA44, BA45 and BA47. Laminar and gyral patterns are generally not tightly mapped on one another and for Broca's area in particular, there is significant intersubject variability (Amunts et al. 1999). However, there is still a loose correspondence between pars opercularis and BA44, pars triangularis and BA45 and pars orbitalis and BA47. The confusion regarding the anatomical borders of the area arises: (a) because Broca's area was initially defined functionally and based on gross anatomical descriptions, (b) because it is frequently described in the literature either in terms of either cytoarchitectonics or gyral anatomy and (c) because it is not yet clear whether all three Brodmann areas and gyral parts should be included in the description of the region. Different studies use different definitions of the area leading to some confusion with respect to the exact localization of the activations and the actual function of the region and its subregions. However, a clear understanding of the anatomy of the region is fundamental, since anatomy can provide insights on the function of an area and the three Brodmann areas are quite different both anatomically and functionally.

The functional and anatomical difference between the different parts has been shown by studies involving both anatomical and functional connectivity. In a functional connectivity study by Bokde et al. (2001), it was shown that the dorsal and ventral parts of the LIFG have different connectivity weights to the posterior part of the superior temporal gyrus as a function of the stimulus' lexical status (word, non-word, false fonts). These results are in agreement with a recent diffusion tensor imaging (DTI) study that showed that the different Brodmann areas have different connectivity patterns with the rest of the cortex (Anwander et al. 2007). Furthermore, according to studies on the neurochemical fingerprinting of the areas (Amunts and Zilles 2006), the density of glutamatergic AMPA receptor binding sites follows a caudal-to-rostral gradient from BA44 to BA45, with higher concentrations of binding

sites in BA44 than BA45. In terms of the laminar structures, the difference between the three regions is particularly pronounced between BA44 and the two prefrontal regions BA45 and BA47. Located on the borders between the prefrontal and the premotor cortex, the three Brodmann areas seem to be a reflection of this transition both functionally and anatomically. In terms of laminar patterns, BA45 and BA47 are very similar to other prefrontal regions, while BA44 seems to have elements of both prefrontal and premotor regions. BA44 is neither agranular (like premotor area BA6) nor granular (like prefrontal area BA45), but rather dysgranular (Amunts and Zilles 2006).

Agranular cortex lacks layers II or IV (or both) and is therefore composed of two or three cellular strata. Granular cortex on the other hand contains distinct granule cell clusters in layers II and IV and is therefore composed of 4 or 5 cellular strata. The dysgranular cortex seems to represent an intermediate stage between the two types of cortices, in which layers II and IV are not clearly distinguished (Mesulam and Mufson 1982). In the case of BA44, layer IV, the layer where granule cells are located, is not well developed and large pyramidal cells from layer III seem to invade the layer. This is in contrast to BA45, where layer IV is easily distinguishable (see Figure 3 for more details). Based on such evidence it was suggested that region BA44 is a transitional area between region BA45 and BA6 and consequently also between the granular prefrontal and the agranular premotor cortex. The functional and anatomical implications of the transitional, prefrontal and premotor character of region BA44 are not yet well understood, but are subjects of current research.

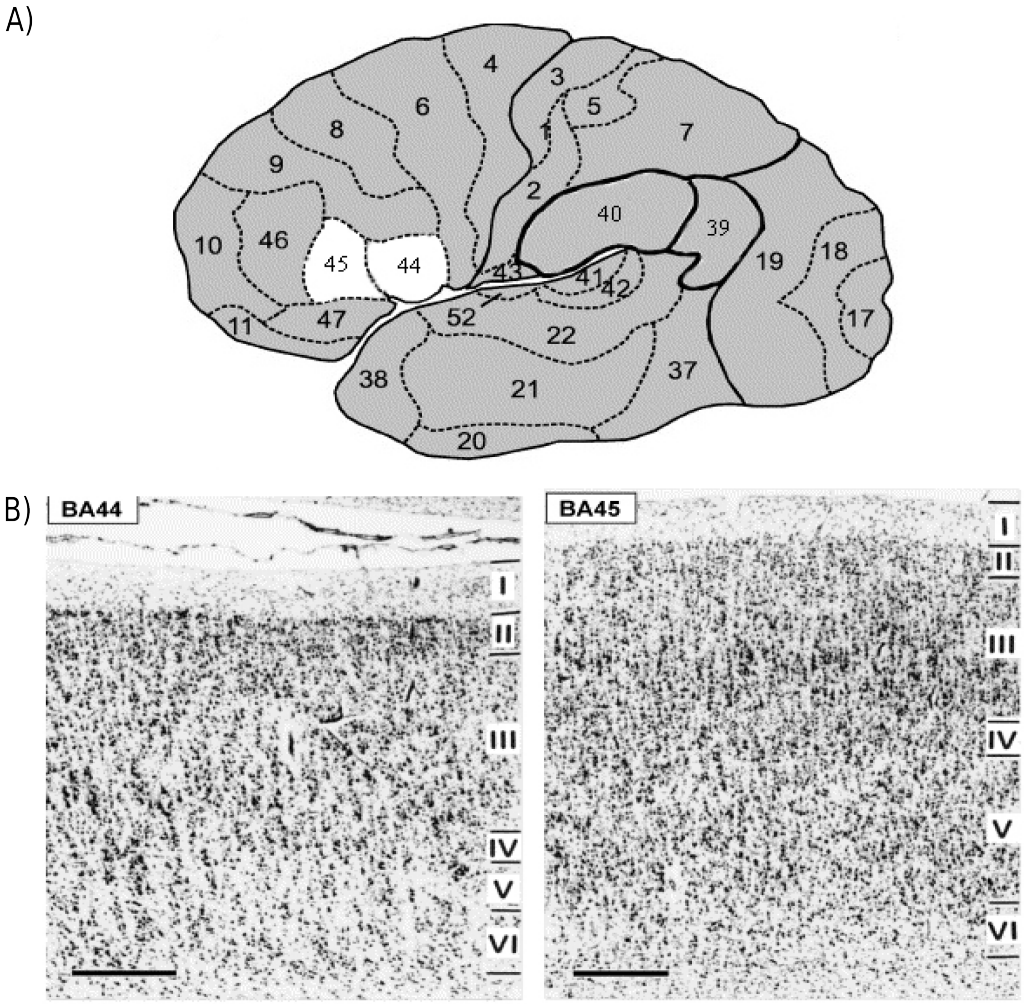

Figure 3: Cytoarchitecture of BA44 and 45. Shown in (A) is a lateral surface of the human cortex divided into different cytoarchitectonic areas following the work of Brodmann (Brodmann 1909). BA44 and 45 are shown in white. Shown in (B) is a coronal, cell-body stained section of a post-mortem brain for areas BA44 (left) and 45 (right). The cytoarchitecture of both areas is characterized by large pyramidal cells in deep layer III, which exceed in size those of layer IV. Whereas granular BA45 shows a clearly visible layer IV, the layer IV of dysgranular BA44 is thinner and not clearly discernible from neighbouring layers, since it is invaded by pyramidal cells from layers III and V. Cortical layers are numbered with Latin numbers. Scale bars are 0.5 mm. Image adapted from Amunts et al. 1999.

1.3.2 The Function of Broca's Area

Based on what has already been mentioned, if one assumes a wide perspective in the anatomical borders of Broca's area, he/she will end up including both prefrontal and premotor subregions. In this sense it is not surprising that activation in this area has been reported for a variety of cognitive tasks. The prefrontal cortex is associated with a large number of high-level cognitive processes such as language, working memory, abstract reasoning, problem solving etc. The premotor cortex on the other hand is mostly associated with motor functions and processes related to motor movement planning. However, new evidence now suggest that it also has a role in higher cognitive processes and is involved in spatial perception and action understanding (Rizzolatti et al. 2002).

As with its anatomy, the function of Broca's area appears to be related both to prefrontal and premotor functions. The results from numerous neurophysiological and neuropsychological studies support the hypothesis that the region's prefrontal properties are associated with language processing (e.g. phonemic structure processing, verbal working memory and sentence planning), while its premotor functions include more general motor planning, imitation and most importantly speech (for a review on the functions of Broca's area see (Nishitani et al. 2005). What is also interesting, is that some of the studies on the function of the LIFG also show a functional segregation between the different parts of the IFG. This functional segregation is loosely correlated with the anatomical segregation. Each of the three parts, BA44, BA45 and BA47 or pars opercularis, triangularis and orbitalis has been associated with a particular level of processing (Devlin et al. 2003). BA44 and the pars opercularis have been associated with phonological processing, BA45 and the pars triangularis with grammatical and syntactic processing and BA47 and the pars orbitalis with semantic processing.

Since the subject of this thesis is the phonological/phonetic system, a narrow definition of Broca's area, focusing on BA44 and the pars opercularis, would be more suitable. Even though the anatomical correspondence between BA44 and the pars opercularis is only an approximation, at the moment it is not possible to non-invasively map cytoarchitectonic areas on gyral anatomy. For the remainder of this thesis the correspondence between BA44 and pars opercularis will be accepted, although with caution. The term Broca's area will be used to refer to BA44 and the pars opercularis. There will not be any explicit distinction between BA44 and pars opercularis, unless otherwise specified. After this clarification with respect to the anatomy of Broca's area, the next point of focus is function. Discussed in the following paragraphs are some of the theories behind the role of Broca's area in language production and the questions that still need to be answered.

1.3.2.1 Language-Related Processing

As it has been mentioned already, Broca's area was originally thought to play a predominant role in speech production (Broca 1861). However, there has been much disagreement among researchers with respect to the exact role that the region plays in speech production. To complicate things even more, it was also noted that Broca's area is not just involved in speech production, but also speech perception. Recent findings, suggest that the networks supporting speech perception and production are largely overlapping (Heim et al. 2003b; Okada and Hickok 2006a). Further questions that then came up where with respect to the common processes that underlie both speech perception and production and whether Broca's area was actually involved in any of these common processes rather than just speech production or speech perception.

A clear step to this direction was made by Riecker et al. (2005) in a study distinguishing between speech planning and speech execution. They showed that Broca's area is not involved in speech execution, in the sense of articulation, but

rather speech planning. Estimation of the functional connectivity between all the regions that showed a significant effect for a contrast between syllable repetition and passive listening revealed two distinct, left lateralized networks involved in speech planning and execution. The first network consisted of the dorsolateral frontal lobe (including Broca's area), the supplementary motor area (SMA), the anterior insula and the superior cerebellum, and was associated with speech programming. The second network consisted of the primary motor area (M1), the thalamus, the basal ganglia (putamen and caudatum) and the inferior cerebellum, and was related to speech execution. Based on such evidence the authors concluded that Broca's region is involved in speech planning. But what type of representations does the region process?

Earlier studies on lexical processing have identified a number of functions for the region. Studies employing phonological processing tasks such as rhyming judgement (Poldrack et al. 2001), syllable counting (Poldrack et al. 1999) and phonemic discrimination (Zatorre et al. 1996), have proposed a role of the posterior LIFG in phonological processing (also see Zatorre et al. 1996; Poldrack et al. 1999; Bokde et al. 2001; Amunts et al. 2004). However, this region is also thought to be involved in verbal working memory and to facilitate thematic role assignment or sub-vocal rehearsal during delayed response tasks (Caplan 2001; Newman et al. 2003; Tagamets et al. 2000). Anatomically, working memory processes were mapped at the more ventral and rostral part of the region (the lower border between BA44 and BA45), while phonological structure processing is believed to take place at the most posterio-dorsal part, near BA6 and the premotor areas (Zatorre et al. 1996; Chein et al. 2002). This difference was shown in an experiment performed by Zatorre et al. (1996), where they contrasted working memory and structure processing by comparing phonemic monitoring and phonemic discrimination tasks. The former task involves working memory and requires the subject to judge whether two words presented one after the other have the same final phoneme. The latter task, phonemic discrimination, is more related to structure processing and segmentation, and the subject is requested to identify a specific phoneme in an auditorily presented word.

This process forces the subject to segment the phonemic structure of the word. Their findings suggested that the posterior-dorsal portion of the LIFG might be involved in general structure building of a sequence, whether it is phonological or syntactic, while the more ventral part was associated with working memory.

Further insights on the function of the LIFG came from studies on non-human primates, in particular on the mechanisms of temporal sequence processing and storage in the macaque monkey. Recent studies have shown that BA44 and 45 have homologues in the primate cortex (Petrides et al. 2005) and new theories appeared suggesting that the nature of motor planning performed by Broca's area is not necessarily specific and limited to language, but is more general and could be responsible for more general action planning. The next section is dedicated to this issue and will provide an overview of the evidence focusing on the relationship between Broca's area and its homologue region in the macaque brain, as well as the implications for non-language specific functions.

1.3.2.2 Beyond Language

The discovery that Broca's area is not specifically involved in linguistic processes, but is also involved in other types of action processing, made a big impression on the scientific community and raised further questions with respect to the overall function of the region. New evidence came primarily from brain imaging studies on action imitation (Iacoboni et al. 1999; Buccino et al. 2001; Buccino et al. 2004; Binkofski and Buccino 2004) that supported the idea of a human mirror neuron hypothesis (MNS; Gallese et al. 1996; Rizzolatti and Craighero 2004). Based on this hypothesis, Broca's region is part of a larger network facilitating motor planning and execution, but also involved in the process of understanding and learning by imitation (for a review see Iacoboni 2005). This system is not specific to humans, but its analogues can also be found in non-human primates. Though a detailed account of the MNS is outside the scope of this report, we feel that a description of some of the features of

the system would be important for overall coherence.

One of the most striking features of the MNS network is that some of its neural populations seem to be active during both action observation and execution. It has therefore been suggested that a population of neurons is responsible for encoding and retaining information on observed actions and then making this information available when one needs to either repeat the same action or mentally replay it for either understanding or learning processes. For example during movement monitoring, sensory input is mapped on the motor system of the observer, facilitating both understanding and learning. Broca's area is thought to to be part of this system, in addition to the inferior parietal lobe (IPL) and the superior temporal sulcus (STS). These three regions form the minimal human MNS (Rizzolatti and Craighero 2004; Iacoboni 2005). Although their precise interaction is yet unspecified and to a large extent task-related, evidence from studies on macaque monkeys has provided many insights about their function.

According to a theoretical model of imitation of visually presented actions (Arbib 2003; Iacoboni 2005), a sensory representation of the observed action is formed at the posterior part of macaque area V5 (analogue of human posterior STS), a site specific for motion detection. It is then forwarded to the caudal and anterior intraparietal sulcus (cIPS and AIP; analogue to human intraparietal sulcus IPS; Sakata et al. 1995) for further object processing. The AIP is thought to have a role in extracting the affordances[3] of the objects and in forwarding the information further to area F5 so that it can make a decision on the action needed (see Figure 4 for a diagram of the processing stream overlaid on a lateral view of the macaque cerebral cortex; Oztop and Arbib 2002). Region F5 (the ventral part of area 6, which is also considered by some to be the human homologue of BA44) is the site where the overall goal of the action will be identified. A rough motor plan of the action to be imitated will be then constructed in cooperation with area F1 (primary motor area)

3) Affordances are the physical properties of an object that determine its function and use (Gibson 1979)

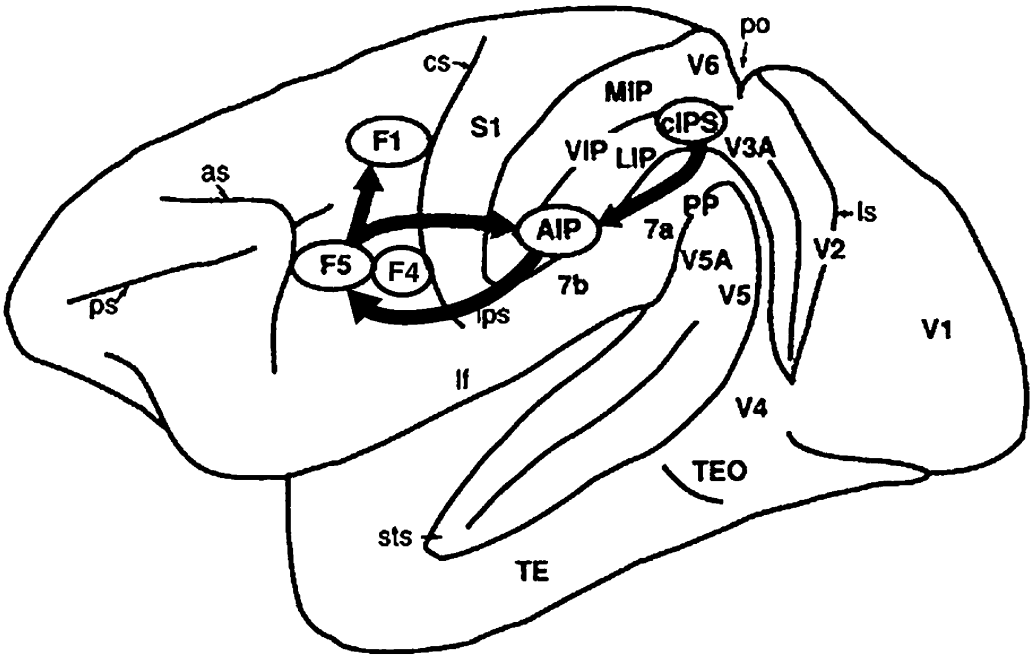

Figure 4: Lateral view of the macaque cerebral cortex. The visuomotor processing stream for grasping movements is indicated by a network of arrows running from the parietal to the frontal lobe. Object features are processed by cIPS and AIP to extract grasp affordances. Information on the affordances are then sent on to the canonical neurons of F5 that choose a particular grasp. The mirror neurons of F5 are active during observation of movement for the purpose of recognition. Finally, the information is passed on to F1 for execution. cIPS, caudal intraparietal sulcus; AIP, anterior intraparietal sulcus. Image adapted from Oztop and Arbib 2002.

and other premotor and SMA regions. In constructing the action plan, the system will focus both on the motion and the object observed, with different subparts of the regions encoding the different types of information. After reconstruction of the motor plan the information stored in area F5 is then sent forward to the primary motor region for execution and back again to V5, via the AIP, in the form of predictions. The main aim of this top-down predictive interaction (F5-V5) is to obtain feedback about the inferences of the system by directly contrasting the predictions of the system to the actual sensory input.

In a few words, a central function of the network proposed so far is sensory-motor mapping, whereby actions observed are mapped on the observer's own motor system. This holds true also for the human neural system, whereby Broca's region is a fundamental part of that network and is in close cooperation with the IPL and STS. According to Iacoboni (2005) its role is to identify the overall goal of an action and construct an appropriate action plan. The exact details of the processes involved still need to be specified. However, this seems to be one of the most complete models of the system that can also account for observations from the field of language processing.

Reported in section 1.2 was a model on language processing that has been proposed by Hickok and Poeppel (2000; 2004). This model has been inspired to a large extent by research in visual processing and the existence of two processing routes, a “what” (ventral pathway) and a “where” (dorsal pathway), hence it has been dubbed the dual-stream model of the functional anatomy of language (DSM). Another source of inspiration for this model was the MNS and the concept of sensory-motor mapping with its extensions to account for linguistic processing. Broca's area, along with regions in the premotor and primary motor cortices and an area in the sylvian parieto-temporal junction, were assigned to the language equivalent of the dorsal “where” pathway (see Figure 5). The “where” pathway is not only responsible for spatial processing, but also for visuo-motor integration (Rizzolatti et al. 1997; Milner and Goodale 1995). For language, this pathway is thought to be responsible for phonological processing and speech production by performing a type of sensory-motor mapping, where the phonological representations are mapped on articulatory representations. This pathway is contrasted to the ventral pathway, where phonological representations are mapped on to conceptual representations. The role of Broca's area is to construct the motor plan for articulation by holding information on the various articulatory codes of different phonetic units, e.g. syllables or phonemes. In this sense, Broca's area is also thought to be the location of the speech sound map.

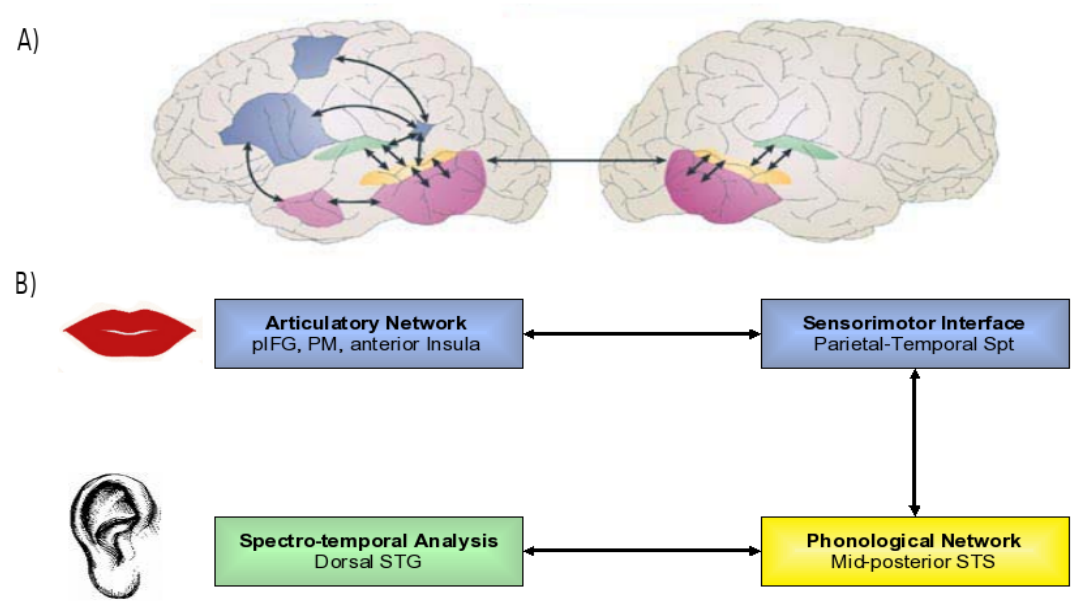

Figure 5: The dual-stream model of language. Shown in (A) is an approximate anatomical map of the dual-stream model components. The earliest stage of speech processing involves some form of spectro-temporal analysis (green), which is carried out in auditory cortices bilaterally. Phonological-level processing and representation involves the middle to posterior portions of the superior temporal sulcus (STS) bilaterally (yellow), although there may be a weak left-hemisphere bias. The dorsal pathway (blue) maps sensory or phonological representations onto articulatory motor representations, while the ventral pathway (pink) maps sensory or phonological representations on to lexical conceptual representations. The posterior region of the dorsal stream, the sylvian parieto-temporal junction (Spt), is proposed to be a sensory-motor interface, whereas the more anterior locations in the frontal lobe, probably involving Broca's region (pIFG) and a more dorsal premotor site (PM), correspond to portions of the articulatory network. Shown in (B) is a schematic depiction of the dorsal pathway highlighting the most important processing components and their associated regions. The same colour-code is used as in (A). Image adapted from Hickok and Poeppel 2007.

Assuming this function of sensory-motor mapping, Hickok and Poeppel then go on to explain functions such as temporal sequence processing and verbal working memory (vWM). For vWM in particular, they argued that it is a special case of auditory-motor integration (Hickok and Poeppel 2000; 2004) and in this sense it can also be viewed as a form of sensory-motor integration (Wilson 2001). One of the dominant models on vWM includes an account of the so called phonological loop

(Baddeley 1992; 2003). In essence, the phonological loop is a mechanism for using a motor system (via articulatory rehearsal) to keep sensory-based (phonological) representations active. It is thought to consist of two components, phonological storage and sub-vocal rehearsal, and it is also thought to be facilitated by the process of generating speech-motor programs, i.e. articulatory codes.

Initial studies on the characteristics of the phonological loop have shown that acoustic or phonological similarity between various targets can have a detrimental effect in the process of retaining the target in working memory (Baddeley 1966). These findings suggested that the information relevant to the phonological loop is acoustic or phonological in nature. However, it was later shown that the key process behind sub-vocal rehearsal is related to the construction of speech-motor plans. This assumption was based on findings from a study of a dyspraxic patient (Caplan and Waters 1995) who could not assemble speech-motor control programs and also showed impaired working memory performance, even though the patient's ability to process phonological information and to articulate was relatively intact. In this case, only the mechanism of speech planning was impaired. However, this impairment also had an effect on the patient's performance in short-term memory tasks, by affecting the patient's rehearsal functions. It was therefore concluded that speech-motor programs underpin sub-vocal rehearsal and verbal working memory.

Such findings are in agreement with some of the claims made by Hickok and Poeppel. However, in the original Baddeley model, a direct translation between phonological and articulatory representations was not proposed or implied. What Hickok and Poeppel point out is that the mechanism of the phonological loop, i.e. the interaction between the phonological store and the sub-vocal rehearsal module, could be greatly facilitated by a direct sensory-motor representation or, at least, such an account is not incompatible with the theory of the phonological loop (Hickok and Poeppel 2004).

Hickok and Poeppel also proceed to specify the role and significance of such a

sensory-motor mapping system. Studies on patients suffering from Broca's motor aphasia have previously shown that even though these patients have an extended portion of the posterior LIFG destroyed and a deteriorated ability to speak, they still retain a good level of understanding (for a review on evidence from neurophysiological studies see Hickok and Poeppel 2004). If Broca's area is indeed facilitating the transcription of sensory information to the primary motor area and if this transcription is essential for successful communication, one would assume that these patients would have severe difficulties in comprehension, which is not the case. Patients that do exhibit severe understanding difficulties are those with lesions in the posterior STS (also known as Wernicke's area). Hickok and Poeppel took this as evidence to suggest that, at least in speech comprehension of adults, sensory-motor mapping only has a secondary role (Hickok and Poeppel 2004). Their view is that the process is particularly important in the early years of life during language learning, or in cases of foreign language learning, when a person is found in a new and more demanding linguistic environment and new articulatory codes need to be compiled. However, it is should not be necessary for everyday communication.

From what has been mentioned so far, it seems like the MNS framework can account for many observations in cognitive research. However, so far only the positive side of the argument has been presented. Naturally, there are also weaknesses in this theory and counter-arguments. One of the weaknesses of the human MNS theory, particularly with respect to the role of Broca's area, is that many of the insights on the role of this area have been borrowed from studies on macaque monkeys and are based on the suggested cytoarchitectonic homology between macaque area F5 and human BA44. A major issue is the fact that this relationship between the two regions has not been proven to satisfaction and there are clear anatomical differences between the two areas. Region F5 is part of lower premotor area 6 and it forms the anterior part of the ventral agranular premotor cortex, i.e. it lacks layer IV (Petrides 2006). In contrast, human region BA44 does not clearly belong to the premotor cortex, but seems to be an intermediate area. Therefore, the search for a BA44 homologue had not been resolved, until very recently. Petrides et al.

(2005) discovered a small area buried within the posterior bank and fundus of the arcuate sulcus, which is dysgranular, exactly like BA44. Because of the location of this area (just anterior to the ventral part of BA6) and its structural characteristics (unformed layer IV), they considered it to be comparable to human BA44. However, since this discovery is very recent, to our knowledge there have not been any published studies on the functional properties of this area and whether it would show the same behaviour as F5.

Continuing with the counter-arguments of the MNS theory, it is not only the anatomical similarity of BA44 and F5 that has been questioned, but also their functional similarity and particularly the role of BA44 in imitation and the MNS. Region F5 is suggested to be sensitive to meaningful, goal-directed actions and it is very strongly activated during conditions of imitation. However, recent studies suggest that BA44 may not serve similar functions (Grezes et al. 2003; Makuuchi 2005). A very interesting characteristic of region F5 is that it consists of two anatomically and functionally distinct neuronal populations, the canonical and the mirror neurons (Gallese et al. 1996; Murata et al. 1997). Both populations are sensitive to goal-directed actions, although canonical neurons seem to encode the affordances of an object. Therefore, they will respond even to the presentation of the potential target, irrespective of whether an agent is interacting with the target or not. Mirror neurons, on the other hand, will respond to the interaction between an agent and an object, irrespective of whether the subject is performing or observing the interaction.

If the two regions, Broca's area and F5 were truly homologue regions, one would expect to observe the same pattern of activation for the LIFG. Even though the lower spatial resolution provided by non-invasive imaging studies on humans (e.g. fMRI) is not suitable to identify neuronal populations as clearly as electrophysiological studies, one would still expect that the region containing the neuronal populations, in this case the LIFG, will show activation patterns that would reflect the presence of both neuronal populations. However, an attempt to map these two populations on the

human cortex using fMRI did not identify Broca's area as the main area of activation, but a more dorsal region on the ventral limb of the premotor cortex (Grezes et al. 2003). Broca's region was sensitive to action-planning with respect to observed objects and imitating gestures, and it also showed significant activation (vs. baseline) during conditions of object, gesture and goal-directed action observation. These results suggested that the area did not make a distinction between canonical and mirror neuron features and did not fit well into the proposed role or at least not as well as the ventral premotor cortex. These results are in agreement with the recent claim that macaque area F5 is actually a homologue of the human ventral BA6 (Petrides et al. 2005).

More evidence against a possible role of Broca's area in imitation was also provided in a recent study contrasting imitation and cued response (Makuuchi 2005). It has been argued that using simple and very familiar actions, like grasping, as tasks for imitation, is not appropriate for adults, who can perform the movement without the need to really imitate it. More variation and complexity in the tasks is therefore needed before it can really be claimed that Broca's area is involved in imitation. Makuuchi (2005) used more demanding actions as stimuli for imitation and by delaying the period between observation and execution, he was able to show that Broca's area is primarily sensitive to delay variation and not imitation. Therefore, it was suggested that Broca's region is related to action planning and working memory related processes and not sensory-motor mapping. Specifically, it was argued that Broca's region acts as a temporary storage of sensory related information that is used for motor preparation.

To a certain extent, this view is similar to a hypothesis developed for the role of Broca's area in lexical processing by Zatorre et al. (1996) and Chein et al. (2002). As previously mentioned, Zatorre et al. have previously shown that Broca's region is sensitive to both working memory demands and temporal sequence structure processing, following a ventral-dorsal segregation (Zatorre et al. 1996). The processing of lexical and non-lexical stimuli could therefore converge to a common

hypothesis on the role of Broca's area and a similar network. In this description of working memory, sensory-motor mapping is not considered to be a facilitatory mechanism and the role of Broca's area would be related specifically to verbal working memory, as has been argued by Baddeley (1992; 2003). This account of working memory does not necessarily require a direct sensory-motor translation, and in this sense conflicts with the theories of Hickok and Poeppel (2004).

To summarize, the function of Broca's area is still as debated as it has ever been. If one follows the claims of Indefrey and Levelt (2000; 2004), Broca's area plays a key role in phonological processing and syllabification. However, Hickok and Poeppel (2000; 2004; 2007) argued that the region's primary role is in sensory-motor mapping and specifically, phonological to articulatory code translation. It should also be noted that neither Indefrey and Levelt nor Hickok and Poeppel included any accounts in their models of a functional segregation within Broca's area. In this sense both of these models are in contrast to much evidence from neuroimaging studies that suggest a functional segregation of the area in its three cytoarchitectonic areas, as well as a dorsal-to-ventral gradient (Devlin et al. 2003; Anwander et al. 2007; Chein et al. 2002; Zatorre et al. 1996). Therefore, questions remain: What is the function of Broca's area? Which theory best accounts for the observed data? The following chapters describe the current efforts to resolve these questions and thereby gain a better understanding of the role of this area in language and phonological processing.

1.4 Experimental Questions and Hypotheses

In this thesis, we[4] investigate the regions involved in the generation of articulatory codes and in particular the role of Broca’s area in the process of generating an articulatory motor plan. Since it has been shown that the posterior part of Broca's area is involved in phonological processing (Devlin et al. 2003), we focused on this

4) Please refer to appendix B for a description of the division of labour among the author of this thesis and other involved parties.

area, which is roughly equivalent to BA44. We specifically wanted to address whether this region is involved in (1) phonological processes, such as syllabification (Indefrey and Levelt 2000), (2) directly retrieving/compiling the articulatory gestures (Hickok and Poeppel 2004) or (3) sub-vocal rehearsal and verbal working memory (Baddeley 2003). We also wanted to examine whether there is a functional segregation in this area between a dorsal and a ventral part. To address these issues, we identified contrasting hypotheses between the different models and designed a series of fMRI experiments to examine the activation patterns resulting from the experimental manipulations and particularly the behaviour of the LIFG.

The first question that we wanted to address is whether Broca's area is involved in phonological or phonetic encoding. The two hypotheses make different predictions about the sensitivity of the region in sub-lexical frequency effects. As previously mentioned, it has been suggested that low and high sub-lexical frequency syllables are processed differently in the brain. High sub-lexical frequency syllables are pre-compiled and stored in the mental syllabary, while low frequency syllables need to be compiled on-line. Phonetic encoding is the mechanism of generating articulatory codes and as a process it is sensitive to the above difference. If the posterior part of Broca’s area is only involved in the process of syllabification, it should not show a significant effect for sub-lexical frequency manipulations. On the other hand, if the area is involved in syllable articulatory code production, we expect the effect to be significant and to observe higher activation for low vs. high sub-lexical frequency syllables in cortical areas that are involved in compiling the articulatory scores.

To address these questions, we used event-related fMRI to monitor the changes in blood oxygenation while subjects performed a delayed pseudoword repetition task. The presented pseudowords were constructed so as to be different in both length (four vs. two syllables) and sub-lexical frequency of components (low vs. high sub-lexical frequency). We hypothesized that by experimental manipulation of stimulus length, the network underlying phonological and phonetic encoding would show higher activation for longer vs. shorter words, since longer targets have longer

processing time and require more processing resources. The resulting network would show the regions underlying the system of phonological and phonetic encoding.

Manipulating sub-lexical frequency allowed the identification of the areas specifically participating in compiling the articulatory codes for given phonological codes. We expected that a subset of the identified network for phonological processing would also show a significant activation for the contrast between low and high sub-lexical frequency stimuli. These regions would comprise the network underlying the generation of articulatory codes. We anticipated the functional contrast low vs. high frequency pseudowords would reveal the regions participating in on-line articulatory code generation, while the contrast high vs. low frequency pseudowords would show the location of the mental syllabary. As previously mentioned, if Broca's area is involved in syllabification and phonological processing prior to the encoding of the articulatory scores, it would only show a strong effect of length, but not frequency. On the other hand, if Broca’s area is the site of the mental syllabary, we expected to see significant effects of both length and frequency manipulations.

A potential confound of the experiment designed above is that the presence of a delay period. Even though the delay period is a constant factor across the conditions of interest (sub-lexical frequency and length), it is possible that the effects that we observed are dependent on the activation of the phonological loop and verbal working memory and not related to phonological processing per se. To address this concern, we performed a second event-related fMRI experiment that did not involve verbal working memory. This experiment was similar to the previous phonological repetition task with the exception of the delay period. There was no delay either between the stimulus and the response probe or between the response probe and the subject response. The presented pseudowords were constructed much like the dataset used in the first experiment using the same biphones, but in different combinations, so that the resulting pseudowords were different, but maintained the same statistical and phonetic characteristics. Once again, pseudowords were different in length and

sub-lexical frequency.

If Broca's area is involved in phonetic or phonological processing independent of demands on working memory, we expected that we would be able to replicate the results of the previous study involving delayed phonological repetition. If the posterior part of Broca's area is involved in the process of phonetic encoding, it should show a significant effect for sub-lexical frequency manipulations during a prompt response task. On the other hand, if the area is not involved in syllable articulatory code production per se, we expect that the effect will not be significant. Based on the theory on the existence of a mental syllabary, we expect that frequently used syllables would be pre-compiled and stored in the area, while infrequent ones would need to be compiled on-line based on their segmental features, i.e. phonemes. If this theory is correct, then we should be able to observe the same effects independent of whether the task involves a delay or not.

During these experiments we also examined whether we could identify a functional segregation within Broca's area, as has been reported elsewhere (Zatorre et al. 1996; Chein et al. 2002; Molnar-Szakacs et al. 2005). To identify whether there is a functional segregation within the area, we observed the anatomical characteristics of the functional activation maps for the different conditions and compared across conditions. We also performed a further series of high spatial resolution fMRI replications, focusing on the LIFG. The purpose of these studies was to provide more evidence about functional segregation within the LIFG and specifically the anatomical details of the segregation. The results of these studies will be presented in the following chapters, following an introduction of the methods of data collection and analysis.

Chapter 2: Data Presentation and Collection Methods

To address the experimental questions discussed above, we designed a series of event-related functional magnetic resonance imaging (fMRI) studies. We used auditory stimuli and manipulated phonological and phonetic properties to create experimental contrasts between different sub-lexical conditions, such as target length and sub-lexical frequency. The technique of fMRI was the most appropriate for our study, because of the good spatial and temporal resolution available. In our study we had a specific hypothesis about the role of the LIFG and we were also interested to see whether we could observe any functional segregation of the region, which, non-invasively, would be possible only through the use of fMRI. In this chapter we will describe the features of the experimental stimuli, the design and the technique used for our series of studies, to gain a better understanding of what we are measuring and how to interpret our results.

2.1 Stimuli

Because in this study we were interesting in studying the phonological system, we chose to use pseudowords instead of words or non-words. We avoided the use of words that exist in the lexicon, because of the potential confound of semantic effects, such as lexical frequency. We also did not use non-words. The difference between pseudowords and non-words is that pseudowords are meant to be made-up words that are phonotactically legal and pronounceable. In our study, we are interested in comparing low vs. high sub-lexical frequency pseudowords and in essence we are comparing between segments (e.g. syllables) that are pre-compiled vs. segments that are compiled on-line. Therefore, it is important that the experimental stimuli are pronounceable and legal, so that the main difference between the contrasting conditions would be the stimulus length and their sub-lexical frequency.

Four sets of 72 pseudowords were created (a total of 288 items) varying in length and sub-lexical frequency: four-syllable low frequency, four-syllable high frequency, two-syllable low frequency and two-syllable high frequency. Half of the stimuli (36) per category were used in the delayed response experiments and the other half in the prompt response experiments. The four sets of stimuli consisted of alternating consonant-vowel (CV) biphones plus a final consonant, i.e. CVCVC and CVCVCVCVC for two and four-syllable pseudowords respectively. The four-syllable pseudowords contained two stresses (a primary and secondary stress). However, the position of the stressed syllables within the pseudowords varied to allow greater flexibility in the creation of the dataset and avoid the creation of ungrammatical syllables. Examples of the stimuli are presented in Table 1 and a full list of the stimuli used can be found in appendix A. As a measure of length we chose number of syllables and phonemes, with minimum stimulus length of two syllables. Two-syllable pseudowords were preferred over monosyllabic ones to allow better control of phonological neighbourhood density, which decreases as the length

Table 1: Stimulus Features

Condition	Biphone PP	Phoneme PP
4 syllables, high PP e.g. \hɛ.tə.tɛ.sɜ˞g\	0.0251 (±0.0093)	0.4888 (±0.0681)
4 syllables, low PP e.g. \gɔ.fɑ.θo^{w}.jɜ˞g\	0.0013 (±0.0012)	0.1251 (±0.025)
2 syllables, high PP e.g. \kɪ.kɛb\	0.0181 (±0.007)	0.2965 (±0.0427)
2 syllables, low PP e.g. \go^{I}.tʃɜ˞z\	0.0004 (±0.0004)	0.061(±0.0194)

Note: Table with examples of the stimuli used in each category (phonetic transcription) and their features. For each category we include the mean (±std) phonotactic probability (PP) measures for both biphones and phonemes.

increases (Pisoni et al. 1985). As a measure of sub-lexical frequency we chose the phonotactic probability of the individual phonemes and biphones. Phonotactic probability refers to the frequency with which legal phonological segments and sequences of segments (e.g. biphones) occur in a given language (Jusczyk et al. 1994). As observed in the syllable-frequency effect, low phonotactic probability pseudowords and non-words have slower response time than high phonotactic probability ones, reflecting the load in the phonetic encoding process (Vitevitch et al. 1997; Vitevitch and Luce 1998; Vitevitch et al. 1999).

All the syllables, with the exception of two, that were used in the study to construct the pseudowords were chosen from a corpus of previous linguistic studies on the effects of phonotactic probability (Vitevitch et al. 1997; Frisch et al. 2000) such that they were rare, but not illegal (in the case of low frequency items), and that they satisfied our criteria for frequency. The two additional syllables that we included

were /θoʷ/ and /θɚ/[5]. Both of these syllables had a biphone probability greater than zero and were included to increase the variability of the generated dataset. The phonotactic probability for each biphone and phoneme was calculated (Vitevitch and Luce 2004) and pseudowords were created such that they consisted entirely of high or low probability segments (depending on the category).

To reduce the amount of similarity between the stimuli, no two syllables occurred in the same word more than once and no pseudoword appeared as a contiguous part within another pseudoword. All items were further checked for immediate phonological neighbours using a "one phoneme change" rule, i.e. no stimulus could be turned into an English word by (1) substituting one phoneme with another, (2) deleting one phoneme or (3) adding one phoneme. Even though phonological neighbourhood density and phonotactic probability are correlated, we expected that by controlling for immediate neighbours, the differences in neighbourhood density between items with different phonotactic probability would not be emphasized. Effects related to phonotactic probability would then be related to phonetic encoding and not phonological word retrieval, which would arise by manipulating phonological neighbourhood density (Okada and Hickok 2006b). As a result, low and high sub-lexical frequency items differed systematically only with respect to the positional frequency of their phonemes and syllables. Finally, to avoid morphological confounds, any sequences that ended in high probability final rimes which could be interpreted as inflectional suffixes, e.g. /-æs/ and /-æd/, were also omitted from the dataset.

To record the stimuli, we recruited a native, female American English speaker. Prior to the recording, the volunteer was trained to pronounce the dataset correctly and rehearsed the items a number of times to familiarize herself with the dataset. The stimuli were read from a laptop screen and spoken in isolation as naturally and as clearly as possible. All stimuli were recorded in a single session in a non-echoic, sound attenuated booth. They were digitally recorded using a Shure SM58 vocal

5) For transcribing spoken stimuli we are using the international phonetic alphabet (IPA; 1999)

microphone at 44.1 kHz sampling rate and were saved at 16-bit resolution. Two or three recordings were made for every stimulus, which were later edited into individual files and screened for both accuracy and fluency. The most accurate recording of each item was chosen for the stimulus list. The chosen stimuli were then transcribed and their segment and biphone phonotactic probably was recalculated to take into account the cases where there were some differences in the pronunciation. In the resulting lists, the differences between the average segment and biphone probabilities over both four and two-syllable pseudowords were statistically significant (phonemes: $F(1,286) = 920.2$, $p < 0.001$; biphones: $F(1, 286) = 763.9$, $p < 0.001$). Higher frequency pseudowords had higher phonotactic probability scores than lower frequency pseudowords (for more details on the category phonotactic probability see Table 1).

2.2 Experimental Design and Procedures

Stimulus presentation was in a pseudo-random, fast event-related fashion and the occurrence of each event was controlled by a binary maximum length shift register sequence, also known as an m-sequence (Benardete and Victor 1994). The primary reason for using m-sequences for the presentation of the stimuli was that they are easy to implement and offer a high degree of orthogonalization and counterbalancing between events. In the next section we will provide a more detailed description about the experimental design and the features of m-sequences.

2.2.1 M-sequences

When conducting event-related fMRI experiments, it is important to use a paradigm that can provide a good estimation of the haemodynamic response function (HRF) for a given condition and thus increases the efficiency of the design. Experimental efficiency depends critically on the temporal arrangement of the sequence of events

and the noise in the fMRI signal. The use of m-sequences provides a simple, but robust way of maximizing efficiency. Essentially, m-sequences are pseudo-random sequences of integers that assume L different values, where for a binary sequence L = 2. They are generated recurrently from linear shift registers using modulo L arithmetic:

$$s_k \equiv \sum_{i=1}^{N} c_i * s_{k-i}$$

s_k is the next member to be appended to the existing sequence, c_i are recurrent coefficients that belong to an N-order shift register and the symbol ≡ denotes congruence (mod L). For binary m-sequences, s_k and c_i assume values of either 0 or 1. For N order registers the length of the sequence is $L^N - 1$. M-sequences are uniquely determined by a set of coefficients c_i and the content of the shift register. The sets of coefficients c_i that produce m-sequences of two, three, and five levels can be found in the literature (Buracas and Boynton 2002).

There are some specific features of m-sequences that help maximize efficiency in the estimation of the HRF. Firstly, the number of event presentations is equal for all event types, which maximizes the number of presentations for all event types (Liu et al. 2001). The only exception is that zero-events are presented n-1 number of times, n being the number of non-zero events presentation. Secondly, m-sequences are nearly orthogonal to cyclically time-shifted versions of themselves. For any phase of a cyclical shift, binary m-sequences of length n deviate from orthogonality only by 1/n (i.e. the autocorrelation of the m-sequence is zero everywhere apart from one bin). All events are therefore being presented an equal number of times, but they are not correlated, making it easy for one to dissociate the effects for each of the events. Finally, the product of two distinct shifts of an m-sequence is a new m-sequence, a third shifted version of the original m-sequence, which is again almost orthogonal to the other two with a 1/n deviation.

This makes binary m-sequences an ideal means for increasing efficiency in the experimental design. The efficiency of the design is maximized when all columns of the design matrix X are orthogonal and X^TX approaches a diagonal matrix (Buracas and Boynton 2002). This condition is met if (a) event vectors for each event type are orthogonal to each other and (b) an event vector is orthogonal to a shifted version of itself. M-sequences satisfy both of these conditions much more closely than average randomly generated sequences.

However, the gain in efficiency comes at a cost of restricted design flexibility, since the constructed event sequences are constrained to certain lengths and numbers of event types. The length of the m-sequence always needs to be some power of the number of presented events minus 1. For a binary m-sequence, as in our case, the possible lengths could have been 1, 3, 7, ..., 63, 127 etc. The number of event types depends on the order of the m-sequence, e.g. two event types for binary sequences. However, in cases where there is a need for more events, there is a possibility to either use more than one cyclical shift of the same binary sequence (which will still be orthogonal to the original sequence) or to create a ternary or a five-level m-sequence.

For this particular study, we used three binary m-sequences. The main sequence used had a length of 63 bins and was used to create two more shifted versions: one shifted by 9 bins and another by 18. The length of each bin, corresponding to a single trial, was 8s. Within this window, a stimulus was initially presented and then followed by a response probe and the subject's response. Two different sounds were used as response probes, a high frequency tone for overt responses and a low frequency tone for the covert responses. The type of stimulus presented (two- or four-syllable and high or low sub-lexical frequency) and the type of response (overt or covert) was determined by the m-sequences. For the m-sequence that controlled the stimulus length, 1 meant four-syllable pseudoword and 0 meant two-syllable; for the sequence that controlled phonotactic probability, 1 meant high and 0 meant low. Finally, for the

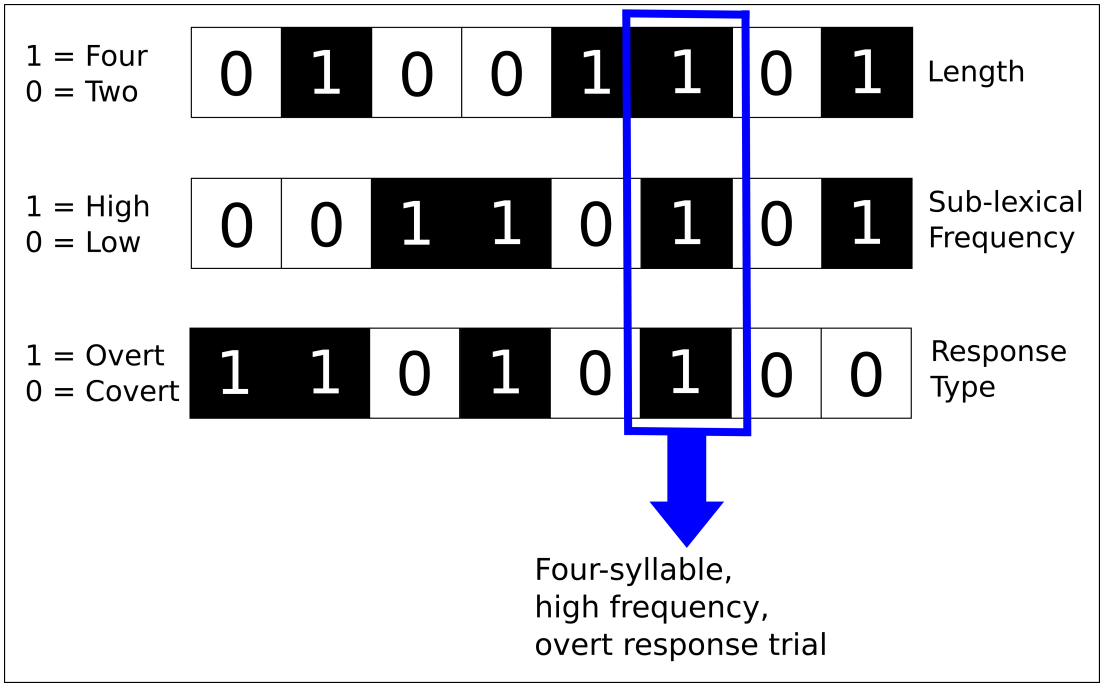

Figure 6: A cartoon of three binary m-sequences, similar to the ones used in the experiment, but shorter for depiction purposes. The two bottom sequences are shifted versions of the top sequence (by 2 and 4 bins respectively). Each bin represents a trial, i.e. in this example there are 8 trials. Each m-sequence controls the presentation of a condition, e.g. the top sequence controls stimulus length. The combination of values across all sequences for each trial determines the type of trial to be presented, e.g. in trial 6 (highlighted with a blue frame) the subject listens to a four-syllable, high sub-lexical frequency pseudoword and is then asked to repeat it overtly.

sequence that controlled response type, 1 meant overt and 0 meant covert response (see Figure 6). Therefore, if all three m-sequences showed 1s for a particular trial, then a four-syllable, high frequency pseudoword would be presented and the subject would be asked to respond overtly (see Figure 6). Because of intrinsic limitations in the experimental design, no null conditions were used.

2.2.2 Scanning Procedures

Two series of experiments were performed, using two different groups of subjects. During both series, a similar scanning preparation protocol was followed to minimize subject movement and ensure comfort. In this section we will describe the basic setup that was common between all scanning sessions. Wherever there were differences between the sessions they will be explicitly described in the respective methods chapter for that study.

As subjects for the experiments, we recruited volunteers from the National Institute on Deafness and Other Communication Disorders (NIDCD) subject pool. All the subjects used in the experimental sessions reported that they were right-handed, American English monolinguals, with normal hearing and with no history of previous neurological or psychiatric disease. Subjects were paid for their participation in the 2-hour scanning session, in compliance with the institutional guidelines. Prior to testing, volunteers provided written informed consent as approved by the NIDCD-NINDS IRB (protocol NIH 92-DC-0178).

In all scanning sessions stimuli were delivered auditorily using an fMRI compatible (pneumatic) system for auditory delivery (Avotec SS-3100, silent scan system). Because the size of the head coil was very narrow, in-the-ear, stethoscopic earphones were used instead of the standard headphones (depicted in Figure 7-B). The tips of these earphones resembled earplugs and were inserted in the subject's ears in the same way as earplugs. They also protected the subject's ears from the scanner noise by offering 30db noise reduction. Prior to the onset of each experimental run and because of the concern that, during the scanning session, the scanner noise would mask out some of the stimuli, a short quality control run was performed. During this run a set of pseudowords[6] was presented to the subjects. The volume of the headset

6) These stimuli were not part of the experimental set, but recorded in the same session as the experimental set, i.e. they had the same amplitude and recording characteristics as the ones used during the actual experiment.

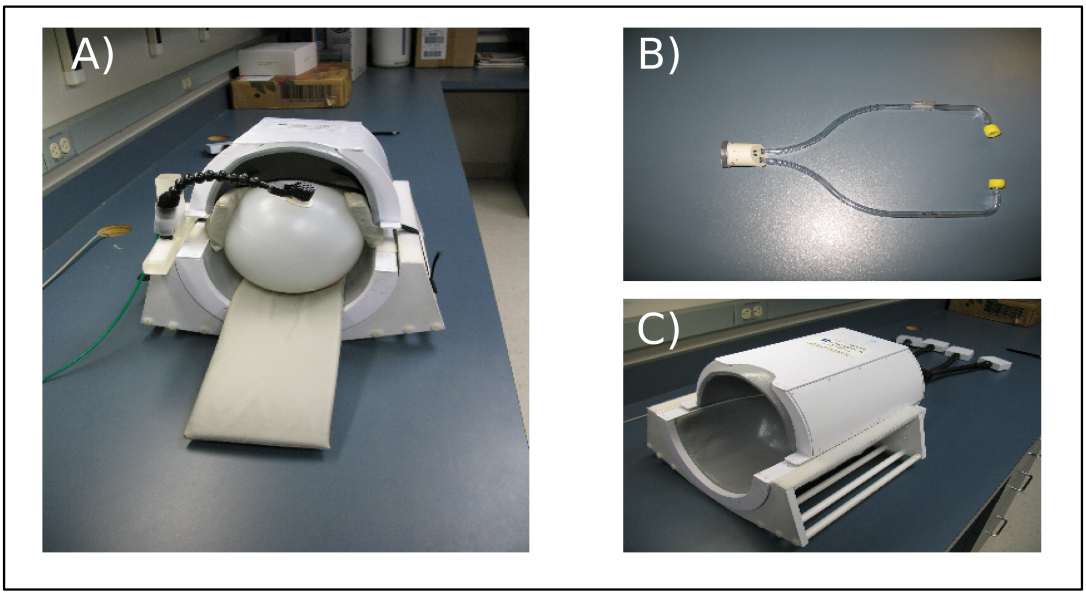

Figure 7: Pictures of the 16-channel coil and the setup used for our fMRI experiments. In (A) a phantom is placed inside the coil surrounded by padding underneath and to the sides. This shows the way that the subject's head was also placed in the coil. A black strap is also tied around the subject's forehead. The MR-compatible microphone is attached to the coil and positioned right in front of the subjects mouth. The stethoscopic earphones are shown in (B). The yellow eartips are inserted into the subject's ear, exactly like earplugs and offer sound protection of 30db. A sideview of the 16-channel head coil is presented in (C).

was then adjusted based on each subject's feedback to ensure protection from exposure to a noisy environment, comfort and clear stimulus delivery. Images acquired during this test run were also submitted to a quality check to make sure that they were free from artifacts. The quality check included a visual inspection of both magnitude and phase images. If there was obvious uneven magnetisation of the head, manual shimming was performed to improve the signal.

During the scanning session subject responses were recorded using a dual-channel, noise cancelling, fibre optic microphone (Dual-Channel Phone-Or by Optoacoustics Ltd., Israel; see Figure 7-A). This system is specifically designed for use in MRI environments and offers real-time adaptive elimination of the MRI acoustic noise from the signal. This allowed us to record both the subject responses and their timing. However, due to concerns that the filtering algorithm introduced a small,

random delay in the recording of the responses, as well as because of the presence of random spikes in the recording of the probe timing, we did not consider the estimates of the subject response timing reliable. Thus, as a behavioural measurement we only used subject response accuracy and the phonotactic probability of the responses.

2.3 Data Collection

For our experiments we chose to use the non-invasive technique of functional magnetic resonance imaging (fMRI). Compared to other non-invasive brain imaging methods, such as electroencephalography (EEG) and magnetoencephalography (MEG), fMRI offers better spatial resolution, though at a cost of temporal resolution. Because we were particularly interested in the anatomical substrates of particular cognitive processes, i.e. phonetic encoding, but also because we had a hypothesis about a specific cortical area, fMRI was the most suitable technique to use. Like every technique, of course it has features and limitations that are important to understand in order to interpret the acquired results correctly. In this section we will describe some of the basic principles behind fMRI, its characteristics and finally the scanning protocol that we employed in our studies.

2.3.1 General Principles of FMRI

A central idea behind fMRI is that neuronal activity requires energy and that the metabolic and vascular processes employed to produce the activity can be visualised. FMRI is a non-invasive, indirect method for measuring and mapping brain activation as a function of cognitive processes. It is non-invasive in the sense that it does not require direct access to the cortex, like other methods such as electrophysiology, nor does it make use of any intravenously applied tracers. Rather it makes use of the magnetic properties of proton nuclei to form a tomograph of the brain that holds

information on the physiological processes that took place at the time of the scan. It is indirect, because it does not record the exact electrochemical processes that are involved in neuronal excitation, but rather measures the vascular changes that arise as a response to brain stimulation. The process of neuronal activation is therefore filtered through the associated vascular and metabolic response (neurovascular coupling) and the resulting map is an image of metabolic and vascular events associated with underlying neuronal events.

This indirect way of measuring the signal means that a lot of information about the neuronal signal is lost and not encoded in the images. For one thing, there is a qualitative reduction of the signal. The different types of brain cell activity (e.g. synaptic excitation, inhibition, action potentials) are reduced to the same signal, since these differences cannot be distinguished in the metabolic/vascular response. Additionally, there is a loss of temporal resolution as the vascular responses are slower than the neural ones and the images produced are on the scale of seconds rather than milliseconds - the scale of the actual neuronal processes. Finally, there is loss of information on spatial resolution, because the recorded vascular processes, usually originating from the arteries and arterioles, are shifted with respect to the actual neuronal processes. This loss is further enhanced by the difference in scale between the spatial resolution of fMRI (millimetres) and the actual resolution of the neuronal processes (micrometres). As a result the fMRI signal is blurred across different types of neuronal populations, which may not necessarily have a similar cognitive function or response to a certain stimulation. In spite all these limitations, the fMRI images succeed in retaining a good spatial resolution on the scale of millimetres, which is still better than other non-invasive methods.

The MRI signal originates in tissue water protons. All nuclei that contain odd numbers of protons, such as the hydrogen nuclei in water, have an intrinsic magnetic moment. In the presence of a strong magnetic field the nuclei tend to assume either a high (oriented against the magnetic field) or a low (aligned to the magnetic field) energy state (see Figure 8-A). To image the location of the resonating nuclei a

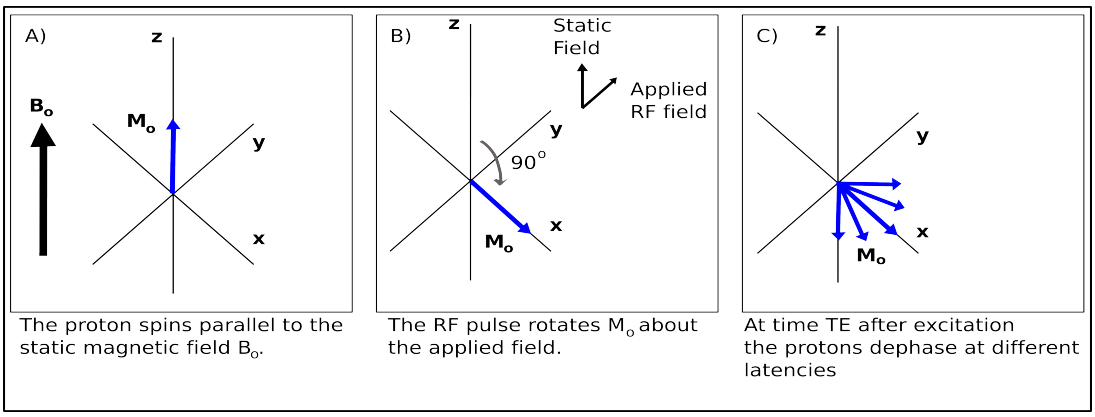

Figure 8: A depiction of the changes in the state of hydrogen protons under different conditions. In (A) a static magnetic field B_o is applied and the protons begin to spin parallel to the orientation of the field. In (B) a gradient pulse is turned on that tilts the direction of the proton spin by 90º. In (C) the gradient has been switched off and after a few milliseconds the proton spins begin to dephase. Different protons will have a different T2 dephase time depending on their surrounding environment. Z is the longitudinal direction and x, y the transverse plane. The cartoon has been adapted after a presentation given by Dr. L. Wald on NMR physics.*

smaller magnetic field gradient (a radio frequency pulse; RF) is superimposed on the larger field. The resonating nuclei at the focus point of the RF pulse will absorb the energy of the pulse and depending on the strength of the RF they will move to a higher energy state, tilting their orientation away from the orientation of the static magnetic field (the higher the strength of the RF pulse, the bigger the tilting angle). The nuclei now 'spin' transverse to the static field and have transitioned from a low-energy state, spinning parallel to the static field, to a high-energy state, spinning at an angle with respect to the static field (see Figure 8-B). After the gradient switches off, the nuclei return under the control of the static magnetic field (see Figure 8-C) and “relax” to their low-energy state by emission of the extra RF energy. This energy can then be encoded by means of a receiver coil into an MR image.

In short, the MR image records the emitted energy from the relaxing nuclei. The differences in relaxation time or proton density between the different types of tissue, e.g. tissue and bones, grey or white matter, forms the basis of the contrast. For fMRI

the source of contrast is related to differences in the NMR relaxation time constants, T1 and T2, of the excited proton nuclei of water molecules. These constants are different depending on the local environment. T1 is the longitudinal magnetisation recovery constant; that is, the time needed for the nuclei to gain full recovery to their original low-energy state and is proportionally correlated to cerebral blood flow (CBF). T2, on the other hand, characterises the signal decay caused by the different spin frequencies that are due to the small differences in the local magnetic environment of each spin (see Figure 8-C). Increase in the T2 value originates from an increase in the blood magnetic susceptibility (BMS) effect, i.e. the extent to which blood modifies the strength of the magnetic field passing through it. Changes in the fMRI signal are therefore due to changes in one or the other of the two relaxation rate constants (Springer et al. 1999), which can be manipulated by adjusting scanning parameters such as the repetition time (RT), the time between successive RF excitation, and the echo time (TE), the delay in encoding after RF excitation.

Depending on the different scanning protocol used, the fMRI signal can depend on different sources. The one that we used in our studies was blood oxygen level dependent (BOLD) fMRI. BOLD fMRI uses the endogenous MRI contrast agent deoxy-haemoglobin (deoxy-Hb) as the source of the contrast. Local neuronal activity, by means of an - as yet unclear - metabolic process[7], induces an increase in CBF and local oxygen delivery to account for the increase in oxygen metabolism (cerebral metabolic rate of oxygen; $CMRO_2$). The coupling, or rather the uncoupling, between oxygen delivery and consumption during elevated neuronal activity forms the basis of BOLD-fMRI (Chen and Ogawa 1999). As has been observed (Fox et al. 1988), the increase in CBF during neuronal activation is much larger than that of $CMRO_2$, which means that there is a surplus of oxygen in the tissue. Oxy-Hb iron atoms are diamagnetic, while deoxy-Hb iron atoms are paramagnetic and increase the local BMS effect, thus reducing the fMRI signal. Therefore, during neuronal activation, the concentration of deoxy-Hb and the local BMS in the activated area

7) A detailed description of this process falls beyond the scope of this study. For a review of the metabolic processes that may be entailed the interested reader is referred to (Magistretti and Pellerin 1999).

decreases, while the fMRI signal increases.

The changes in oxygenation levels can be encoded in the image by means of the T2 relaxation time constant. T2 decay varies exponentially with the levels of oxygen, such that an increase in the concentration of oxy-Hb causes a faster decay rate (decrease in T2). By means of the TE scanning parameter, one can use the differences in decay rate to separate between activated and non-activated areas. Activated areas would have a faster decay rate and higher fMRI signal, than non-activated areas. Typically, the TE for BOLD-fMRI is 20-40ms, which allows for a few percent of BOLD signal change at the tissue.

The BOLD contrast has been used extensively due to its relatively good spatial and temporal resolution. It does however have limitations, which are mainly related to the temporal resolution. In fMRI the temporal resolution is not only limited by the vascular processes, but also by the scanning parameters, i.e. both by the TE and the image acquisition time needed to cover the image k-space. Vascular events occur on a time scale of seconds rather than ms, which is the time scale of the neuronal events. This fact not only sets a limit to the possible temporal resolution of all functional neuroimaging approaches based on vascular coupling, but also introduces some complications related to the design and execution of experiments recording neuro-vascular events. One such complication is related to the minimum duration of a stimulus. Theoretically, even a millisecond-long neuronal event will induce vascular changes. However, in practice it depends on the signal to noise ratio (SNR) whether this change can be measured. Savoy et al. (1995) showed that visual stimulation even as brief as 34ms in duration could elicit small, but detectable signal changes. A further complication, which mainly affects rapid event-related experimental designs, is introduced due to the fact that the haemodynamic response to neuronal events is temporally extended and there is the potential of a non-linear summation of previous and current activations. Despite the fact that non-linearities start to appear even when the intertrial interval is between 2-5s, Dale and Buckner (1997) obtained robust activation for rapidly presented trials spaced as close as 2s. With the introduction of

rapid event-related designs, it was then possible for fMRI experiments to have a more flexible design.

Spatially, the BOLD signal encodes the site of activation by encoding the corresponding vascular changes (CBF and BMS changes). The multiple contributions, however, “blur” the signal, even though differences in $CMRO_2$ and CBF are quite specific to the site of activation. Optical imaging techniques have shown that blood vessels during stimulation become highly oxygenated over an area of a few millimetres in diameter around the site of neuronal activity. This fact determines the intrinsic spatial resolution limit for fMRI. A further limit is also set by the smallest vascular unit that adapts independently to brain activity. Theoretically this is a single capillary, but which source will dominate will depend on the magnetic field. For 3T scanners, the smallest vascular unit is typically the feeding arteriole (> $1mm^3$; Villringer 1999). Magnetic fields greater than 3T are able to record signal from the capillary bed (Chen and Ogawa 1999). However, in magnetic fields less than 3T, a very common magnetic field strength in fMRI research, the signal from the arterioles is also confounded by much stronger activations in the draining veins, which could mean that the activation is displaced by a few millimetres (Lai et al. 1999).

Despite these limitations, fMRI still offers the best spatial resolution for a non-invasive imaging method. In the recent years there have also been several technological advancements to improve both the temporal and the spatial resolution, with the introduction of faster gradient coils, multi-channel coils, parallel imaging protocols and other technological improvements. In our series of experiments, we have made use of recent advances in available technology as much as possible, to improve both the spatial resolution and the signal-to-noise ratio of our data.

2.3.2 Improving Standard FMRI

To counteract the limitations already mentioned and to improve the recorded signal, as well as the signal-to-noise ratio, a number of different approaches have been employed. Here we will only focus on the ones that are particularly relevant for our study.

Noise is one of the worst “enemies” of the fMRI signal. It can arise from physical sources, sometimes referred to as scanner drift (e.g., slowly-varying changes in ambient temperature), from physiological biorhythms (e.g. such as ~1Hz respiratory or ~0.25Hz cardiac cycles aliased by the slower sampling rate) or residual movement artifacts and their interaction with the static magnetic field. Such noise usually appears as a low frequency component of the fMRI time series (Henson 2003). When the subject is performing a task, signal components are also added, which, however, we would wish to distinguish from the noise. Two techniques can be used to counter noise. For one thing, noise can be modelled and the estimated contribution can be removed from the measured signal. Secondly, the experiment can be designed in such a way as to take account of the characteristic features of the system measured and exclude noise from the measurement. “Chopping” for example is a method used to alternate between stimulus or task conditions to generate task-dependent activation with a frequency high enough to minimise noise contributions.

The most common way of eliminating physiological noise or other confounds from various sources is to apply a high-pass filter to the data. With this method, one can remove low-frequency confounds without having to estimate them explicitly. This is also the approach that we used in our studies. We used a cut-off filter of 128s (~0.008Hz) to remove slow signal drifts with a period longer than 128s, which after inspection of the data was determined to be suitable and was not removing much of the event variance.

A further issue that comes up in fMRI is the fact that the head magnetises unevenly because of anatomical differences in magnetic susceptibility and also because of the fact that it is not a perfect sphere. Therefore, there are two main issues here: (a) the presence of an external object, such as the head, will create inhomogeneities in the magnetic field and (b) the heterogeneity of the cortex, due to the presence of materials with different magnetic susceptibility properties, e.g. water, air and bone, will also create distortions and signal loss particularly in the vicinity of the interface between these materials (Buxton 2002). The largest field distortions are due to the air/water interface near the sinus cavities. The presence of these distortions is particularly problematic in echo-planar imaging (EPI), which is commonly used in BOLD-fMRI and can also mean that the signal is spatially displaced. A way to overcome this is to use shim coils, which are used to adjust the magnetic field and correct the non uniformities of the magnet itself, as well as the inhomogeneities of the human head. The geometrical distortions also increase with magnetic field and can be particularly severe for > 3T magnetic fields. If after image acquisition, there are obvious geometrical distortions present, one can correct the distortions during image pre-processing by applying “unwarping” schemes.

Another method to reduce the sensitivity of BOLD-fMRI to geometric distortions is to use sensitivity-encoded (SENSE) echo-planar imaging (de Zwart et al. 2002). This method does not replace shimming, but complements it. For our experiments we used a combination of SENSE-EPI and a 16-channel array coil for additional SNR increase (de Zwart et al. 2004). In short, SENSE allows the single-shot EPI image acquisition duration to be shortened, when compared to conventional, full k-space EPI acquisition. Instead of acquiring the full k-space, parallel imaging techniques use multiple channels in detector arrays to acquire only a portion of the k-space (50% for rate-2 SENSE EPI) and thus achieve undersampling. For a given resolution, the undersampling can be used to reduce image artifacts by shortening the data acquisition window and thus improving image quality (Bammer et al. 2001). For the same image acquisition time this can also lead to the acquisition of more image slices or the acquisition of thinner than usual slices. The latter also adds to the

improvement in image quality. Even though SNR decreases with the size of the voxels, in the case of BOLD-fMRI and single-shot EPI, thinner slices greatly improve SNR by reducing the contribution of physiological noise to the voxel. Simply put, a smaller voxel is more homogeneous than a larger one, which increases the fMRI signal. However, the benefits of SENSE-EPI come with a cost and there is a substantial loss in SNR. The usage of multi-channel receiver arrays can counterbalance that and in some cases (depending on the number of coils, the strength of the magnetic field etc.), it offers an additional increase in SNR, when compared to conventional, full k-space EPI.

2.3.3 Image Sampling Rate

The quality of the data in auditory fMRI can be further affected by the presence of scanner noise, which is created by the switching of the gradient coils every time the MR signal is read out. This noise creates constant activation of the auditory cortex and can also mask the presented auditory stimuli. A solution to this problem is the use of sparse temporal acquisition (STA; Hall et al. 1999) and the acquisition of a single or a cluster of volumes (clustered sparse temporal acquisition - CTA; Zaehle et al. 2007) after stimulus presentation. Because of the filtering of the neuronal signal with the haemodynamic response function (HRF), it is possible to delay the image acquisition to the end of the stimulus and near the maxima and minima of the haemodynamic response. In this way, the effective auditory stimulus for the activation is not masked by the scanner noise, the auditory activation is unaffected by scanner noise and it also enables clear and accurate recording of the subject responses and response time.

However, these benefits are not without a cost in number of samples acquired per trial and a long repetition time (or intercluster interval in the case of CTA) in image acquisition. The first disadvantage affects statistical power, but can be overcome to an extent in CTA. The second disadvantage, the long repetition time in image

acquisition is more significant and puts a constraint on the experimental design that can be implemented. Trials need to be sufficiently long to allow for both presentation of the stimulus in silence and a period of image acquisition long enough to acquire at least one image (or more in the case of CTA).

In our studies, continuous sampling was preferred over sparse sampling, despite the advantages of sparse sampling during auditory tasks. The main reason was that in some of our experimental tasks there were multiple temporal components present (stimulus presentation and response) and if we were to adopt a sparse sampling approach, the length of each trial would be substantially longer. Considering that we were also bound by the length of the m-sequences and the number of event presentation was fixed, this would also mean that the scanning time per run would be substantially longer. Instead we chose continuous sampling, which allowed us to keep the trial length quite short (8s) and the each scanning run to about 9m long. Longer runs would not be recommended as they would increase subject discomfort and reduce attention.

Furthermore, even though our experiment involved auditory stimulation, we were interested only in phonological and phonetic aspects of auditory processing and did not expect these areas to be activated by scanner noise. The major concern was whether the stimuli would be masked by scanner noise. Different imaging protocols, with different noise frequencies and decibel levels, were tested during pilot studies in the scanner room and the one that caused least masking of the stimuli was chosen. During the actual experiment, the auditory delivery system was also adapted for each subject. Finally, recording and analysis of the subject responses during the scanning verified that the subjects perceived the stimulus differences of interest.

2.3.4 Scanning Protocol

For our series of studies, imaging was performed on a 3.0T MRI system (General Electric, Milwaukee, WI, USA), equipped with CRM (Cardiac Resonance Module) whole body gradients. For improved signal-to-noise ratio (SNR) and higher spatial resolution, we used a custom-built 16-channel MRI receive array (Nova Medical, Wilmington, MA; de Zwart et al. 2004) connected to a custom-built 16-channel MRI receiver (an image of the head coil is shown in Figure 7-C). For the functional scans, we used continuous sampling and single-shot rate-2 SENSE EPI (de Zwart et al. 2002). The exact scanning parameters differed slightly between some of the studies and are mentioned in more detail in the methods section of each study. However, as a general rule, four volumes were acquired during each trial. The combination of the dedicated receive array with SENSE EPI allowed a 2- to 4-fold improvement in SNR and a 50% reduction in geometric distortions relative to a conventional setup with a birdcage head coil (de Zwart et al. 2004). As previously mentioned, the reduced geometrical distortions of SENSE EPI were due to the use of a shortened data acquisition window, which also allowed the acquisition of thinner than usual slices.

To increase the efficiency of subject motion correction, for all studies we also acquired isotropic voxels. However, the resulting smaller-than-usual thickness of the slices put a constraint on the brain volume that could be imaged. We were therefore not able to image the whole brain and the size volume imaged depended on the slice thickness chosen in each study. Precise details of the area scanned are mentioned in the methods section for each study. Since all of our studies involved speech, we avoided imaging of the lower parts of the cortex, e.g. the inferior temporal areas, to avoid geometrical distortions and artifacts that are caused by articulatory muscle movement (Birn et al. 2004). To facilitate slice selection, a sagittal two-dimensional anatomical image was acquired prior to the onset of the functional runs. This image was inspected for specific anatomical landmarks such as the anterior commissure and was used to make the slice selection. At the end of the scanning session, high-

resolution spin-echo T1 anatomical images were acquired at the same location as the functional EPI scans. The details of the scanning parameters for the anatomical image for each study can be found in the respective methods section.

To restrain head movement during the scanning sessions, we used head padding and a velcro strap, mounted on each side of the head coil and positioned on the subject's forehead at the line just above the eyebrows (see Figure 7-A). The purpose of the strap was to act as a motion reference point for the subject. Head movement, especially in the z direction, would put a strain on the strap and cause it to rub on the subject's forehead, making them aware of the movement and causing them to restrict it and return to the original position. Prior to the onset of the scanning session the subjects were given instructions about how to restrict their head-movement and about the function of the velcro strap. Tests were also performed to ensure that the strap was properly placed and the subjects could feel it when moving during speech.

Chapter 3: Data Analysis Methods

In the following chapter we will describe the basic principles behind the methods that we used for data analysis. We will first talk about the analysis of the behavioural data and describe both the methods and some of the challenges that we faced. Then we will move on to cover the functional imaging data and describe the preprocessing and statistical analysis approach that we chose. Before the fMRI data can be used in a statistical model, a number of preprocessing steps will need to be performed in order to prepare the data for group analysis and statistical comparisons. Some of these steps include removing artifacts, but also aligning the data from individual subjects to the same space so that they can be used in a group analysis. In the fMRI studies presented in this thesis, we followed a similar preprocessing and analysis protocol. In this chapter we will present an outline of this general protocol, the common methods employed and the basic, underlying principles. Any preprocessing differences or analyses that are specific to an experiment are described in the methods section of the relevant studies.

3.1 Behavioural Data

As part of the experiments performed, subjects were asked to listen to presented pseudowords and repeat them either overtly or covertly. Because half of the responses were covert, we were only able to acquire behavioural measures for half of the presented stimuli. However, because the conditions were all counterbalanced and randomised, we expected that the behavioural data that we would collect from the overt responses would be sufficient to provide us with a representative measurement of subject performance. Behavioural measurements were important for our study, because we wanted to make sure that (a) the subjects were performing the task as instructed and (b) the perception of the stimuli was not disrupted by the scanner noise.

To assess these factors, one of the measures that we used was subject response accuracy. To calculate it, we monitored and phonologically transcribed all recorded subject responses. Because of the low quality of the recording, resulting from the noise reduction filtering, a precise phonetic transcription of the subject response was not always possible and the nearest phonological transcription was used. Cases where the recording was unintelligible because of noise were not included in the analysis. The resulting transcriptions were compared to the target stimulus phoneme-by-phoneme and a score was calculated based on the number of correctly identified phonemes (token count). If a phoneme was omitted in the subject response, it was scored as a mismatch, e.g. if the target was /kɪkeb/ and the response was /keb/, the first two phonemes were counted as a mismatch and the final phonemes were counted as a match. To determine a match between the target and the response we used broad, phonemic criteria and ignored differences between allophones (Vitevitch and Luce 2005). The scores were then submitted to a 2-way ANOVA with factors length and frequency.

Even though we were not able to extract a very detailed phonetic transcription, our interpretation of the data is not dependent on the subtle phonetic details of the subject's performance, e.g. distinguishing between two allophones. Because we were concerned about the fact that the scanner noise would not allow subjects to perceive subtle, between categories differences, such as the use of high or low sub-lexical frequency allophones, the stimuli were generated such that the differences in the phonotactic probability could also be reflected at the phonological level, i.e. the different phones used also corresponded to different phonemes.

On a further note, the primary reasons for analysing the behavioural results were to identify incorrect trials, to ensure that the subjects were performing the task as instructed and that the difference between low and high sub-lexical frequency items was retained in the subject response. For this purpose we also estimated the phonotactic probability of the subject overt responses in the same way as we did for the stimuli (for more details see section 2.1). To determine whether there is a significant difference between the two conditions, we performed t-tests. Finally, we also examined the subject recording to identify trials that were incorrectly answered (i.e. responses on covert trials or no response on overt trials). These trials were excluded from the fMRI data analysis.

3.2 FMRI Data

3.2.1 Preprocessing

In this section we will provide more information on the preprocessing of the fMRI data. This is a very important step, especially in our case. There are two main factors that could potentially affect the quality of our data: the small voxel size, which leads to a decrease in the signal to noise ratio (SNR) per voxel, and subject motion during the overt response condition.

With respect to the SNR, we overcame this problem at the level of image acquisition with the use of multi-channel coils. Under conditions of zero physiological noise, a decrease in the voxel size leads to a decrease in SNR. However, a smaller voxel, especially in the direction of image acquisition (i.e. thinner slices) also suggests a decrease on the impact of physiological noise. Therefore, the two factors counterbalance one another and at the same time the use of a multi-channel coil offers additional SNR increase (de Zwart et al. 2002; 2004). For this study, our main worry is then subject motion, which we tried to overcome both at that level of image acquisition by limiting head movement and at the level of image preprocessing. In the following paragraphs we will be describing the preprocessing steps that we took to prepare the data for statistical comparisons and to correct for head movement related artifacts. All image preprocessing was carried out using the SPM5 software package and associated toolboxes (http://www.fil.ion.ucl.ac.uk/spm/ software/spm5).

The SENSE-EPI images were first reconstructed and transformed into k-space. Each run consisted of 63 trials that were part of an m-sequence, plus 9 more trials that were inserted in the beginning. For each run these first 9 trials (36 images) were discarded from the analysis, because they did not belong to the m-sequence and they would disturb the orthogonality of the conditions. The purpose of adding them to the beginning of each run was to allow the subjects to get used to the task in the scanner environment and for their behaviour to stabilize. Preprocessing further included manually setting the origin of every image (including the anatomical images) to the anterior commissure. After that, images were submitted to slice-timing correction and an optimized motion correction routine to ensure good quality registration (Oakes et al. 2005). For each subject, the functional images were then registered to the respective anatomical image, which had previously been registered to the Montreal Neurological Institute (MNI) anatomical template (based on the icbm_avg_152_t1_tal_lin.mnc template). In the final step the data were transformed into MNI stereotactic space to allow for group comparisons and smoothed with an isotropic Gaussian filter kernel of 6mm (full-width at half maximum) to improve

SNR.

Because we did not acquire functional images of the whole brain, the automatic routine for registering the images with the MNI anatomical template occasionally failed. For each subject, we checked thoroughly the alignment of the anatomical image to the template and if the images were evidently misaligned, we adjusted the orientation of the images manually. During this process we made sure that major anatomical landmarks (both cortical and sub-cortical) were aligned to one another as best as possible. This was a very labour-intensive and time-consuming process. However, it was a crucial step to ensure that the transformation into stereotactic space would not fail. After it was ensured that all subjects were properly aligned with their anatomical images, we proceeded to white-matter segmentation. This step was included to provide priors for the normalization of the images to the MNI anatomical template.

3.2.2 Head Motion Correction

As previously mentioned, head motion causes artifacts and reduces the quality of our data. Therefore, special care was given for the correction of motion related artifacts. In previous studies on speech, it has been shown that subject motion can lead to geometrical distortions in the image and an increase in false positives (Birn et al. 2004). The movement of the head inside the MRI magnet causes some areas, especially those around the edge of the brain, to move in and out of the imaging field of view and to become unevenly magnetised. This is recorded in the image as a change in the signal and when it is correlated with the task (task-related movement, e.g. when the subject is asked to respond overtly) it can increase the number of false positives. In the case of overt speech, false positives tend to appear around the edge of the brain, while geometrical artifacts tend to be more pronounced at the lower parts of the cortex, near the oral cavity. To avoid some of these confounds we did not image the lower parts of the cortex, e.g. the inferior temporal areas, and used thinner

than usual slices that reduce the impact of physiological noise during acquisition. A visual inspection of the images did not reveal any geometrical distortions.

To quantify the effect of subject movement on the quality of our data, we inspected the data from all scanning sessions using the ArtRepair toolbox for SPM5 (Mazaika et al. 2007) and examined the realignment parameters provided by the SPM5 motion correction procedure. The realignment parameters represent the subject head displacement in terms of 6 rigid-body transformations (in 3 translations and 3 rotations). In terms of translations, x is movement along the sagittal plane, y along the coronal and z along the axial. In terms of rotations, roll is movement about the longitudinal axis, yaw about the vertical axis and pitch about the axis perpendicular to the longitudinal plane.

We were particularly interested in scan-to-scan (incremental) motion during the task, i.e. the change in position between the image acquired during the subject response and its immediately preceding image. In previous studies on speech-related motion (Barch et al. 1999), it was found that speech-related motion is mainly scan-to-scan motion primarily affecting the first scan acquired after the response probe. To assess the effects of speech-related motion on our experiments, we performed a three factor ANOVA with within-subject factors response type, stimulus length and sub-lexical frequency, and dependent variables the six motion estimates for incremental (scan-to-scan) movement. The results from this analysis are presented in the methods section of each experiment. In agreement with other studies (Barch et al. 1999; Shuster and Lemieux 2005), the incremental movement was overall quite small and greater for overt response trials than covert response ones. Sub-lexical frequency also had an effect on subject head movement with low frequency items causing greater movement than high frequency items. Because of the significant effects and in order to remove as much of the confounding effects as possible, we also included the realignment parameters in the design matrix as effects of no interest.

Finally, we inspected the movement parameters for extreme movement. We took into

account both incremental movement and absolute movement (i.e. the displacement of a scan with respect to the realignment reference scan of the timeseries, which in our case is the first image in the series). Our criteria for inclusion in the analysis were that a subject would not show absolute motion greater than the voxel size and incremental motion greater than 1mm in translations and 1º in rotations.

Further examination using the ArtRepair toolbox revealed that in a few cases incremental movement even as low as 0.5mm induced global signal changes greater than 1.5% of the mean and “stripe-like” artifacts on the image. To ensure the quality of our data and to completely remove their effect from the analysis we also included an additional regressor in the design matrix for images that showed changes in the global signal greater than 1.5% of the mean followed by a greater than 0.5mm incremental movement (Mazaika et al. 2007).

3.2.3 Analysis

After finishing with the preprocessing, the data were submitted to statistical analyses. There are many methods that one can use to analyse functional data, the most commonly used being a linear regression analysis. An added complication in analysing fMRI data is the fact that the relationship between the stimulus function and the recorded signal is filtered through the haemodynamic response function (HRF), which needs to be modelled explicitly. In the following sections, we will briefly describe the theory behind fMRI analysis and the approach that we have followed in the analyses of our experiments.

3.2.3.1 Linear Regression

Statistical analysis of the factorial event-related experiments was performed using SPM5. The approach followed to estimate the significant effect of an experimental

factor on the dependent variable (the fMRI signal) is an implementation of the general linear model (GLM). In brief, the model used is:

$$Y = \beta X + E \quad (1)$$

where Y is a matrix with information on the observed data, i.e. the BOLD-fMRI signal as reflected in the image signal intensity values. X is the design matrix with information on the timing of events (onsets and durations) convolved with the HRF and other parameters that could describe the signal such as information about physiological noise, subject movement parameters etc. The parameter estimates β describe the contribution of each design factor to the signal and are calculated using weighted least squares (WLS). When estimating the model, we also need to take account of serial correlations that can arise as a result of low frequency noise (biorhythms) and the latency of the HRF (Friston et al. 2000). SPM5 uses an autoregressive model (AR(1)) to calculate the correlations and uses these estimates to correct for non-sphericity during inference by adjusting the statistics and degrees of freedom. Therefore, WLS is implemented by pre-whitening the data and the design matrix with "unbiased" (after estimation of serial correlations) estimates of the error covariance and then using ordinary least squares (OLS). E is the remaining error in the fit.

Because some of our studies involved a delay period and the trials had multiple temporal components, we used a finite-impulse response function (FIR) to model each trial. This approach allowed more flexibility on which components of the trial to model. FIR models are equivalent to selective averaging (Henson 2003), whereby a trial is divided into a number of bins determined by the window length and the duration of the bins. In our case, each trial was modelled using an FIR with 12 bins of 2s duration. As mentioned in section 2.3.1, in rapid event-related designs such as the one we used, there is not enough distance between trials to allow for the BOLD signal to return to baseline before the next trial begins. As a result, we cannot get an independent estimate for each trial and the parameter estimates for each experimental

factor are determined based on the average of all related trials. After the model has been estimated, we end up with 12 parameter estimates per experimental factor.

In more detail, we performed 3-way, random-effects, within-subject ANOVA with factors length (four- vs. two-syllable pseudowords), sub-lexical frequency (low vs. high) and response type (overt vs. covert). Each of the 8 different resulting types of trials (e.g. four-syllable, low frequency, overt response) was modelled by separate regressors and the main effects and interactions were evaluated by contrasting within or across (interactions) the levels of each factor. To perform group statistics we computed the contrast images for each of the 12 FIR regressors per factor. The resulting contrast images from all subjects were submitted to 1-way ANOVA with 12 levels. T-contrasts testing for the predicted shape of the HRF (a canonical, 2 gamma function; (Friston et al. 1998) were performed to produce maximum intensity projections (MIP) and reveal voxels whose differential activity pattern conforms to the shape of the HRF. For the studies that included delayed response, two HRFs were used, one to model stimulus presentation and delay and another one to model the response period. The latter was delayed by 6s relative to stimulus onset, modelling the presentation of the response probe, and it was used to test for significant effects during the response type condition. The studies that involved prompt responses, only used one HRF to model stimulus presentation and response. The response type condition was used as a localizer to allow us to define an independent region of interest (ROI) within the left inferior frontal gyrus (LIFG). Statistical parametric maps (SPM) were thresholded at $p < 0.001$ uncorrected at the voxel level and $p < 0.05$ corrected for family-wise error (FWE) at the cluster level (Hayasaka and Nichols 2003).

Because the use of the FIR includes the danger of fitting noise and increase the number of false positives because of over-fitting, we took a further step to ensure that the significant activations observed were not related to subject motion. We extracted and inspected the parameter estimates for each significantly activated cluster over the window of the FIR (24s). The time course of movement-related activations is

different from that of BOLD related activations. While motion-related signal changes appear as large spikes in the signal intensity and are time-locked to the time of the subject movement, BOLD-related signal changes follow a curve similar to the HRF (Birn et al. 1999). Following neuronal activation, the BOLD signal in the human auditory cortex peaks 4-8s after stimulation onset and reaches 10% of the baseline 5-9s after stimulus cessation (Belin et al. 1999; Hall et al. 1999). It is therefore easy to distinguish between motion and BOLD-related signals. It should also be noted that significant effects for length and frequency were estimated over both covert and overt responses and so we expected that the contribution of motion related artifacts to the significant activations observed would not be as strong. Finally, it has also been shown that in group-level results, the presence of significant motion related effects is minimal, since the site of the motion related artifacts is different across subjects (Barch et al. 1999).

Previously, we presented the methods for performing a whole-brain analysis. However, in some cases we were interested in looking at the effects of one specific region, i.e. the LIFG. The main reason behind limiting the search volume in an fMRI analysis is to bypass the multiple comparisons problem. The fMRI analysis approach that we have described so far is a mass univariate approach, where every voxel in the timeseries is submitted to a statistical test. As a result, for a given statistical comparison there are as many t-tests performed as there are voxels in an image. However, when the number of comparisons increases, so does the number of false positives. To constrain the level of type I error one has to correct for the number of comparisons. For the whole-brain analysis, we have applied a family-wise error correction (FWE), which takes into account the spatial smoothness of the images. Another way of dealing with the multiple comparisons problem is to constrain the search volume used for the statistical analysis. If there is an a-priori hypothesis about the behaviour of a cortical area, the statistical analysis can be limited to this area. In this way the number of comparisons can be effectively reduced.

There are two approaches that one can follow to look at the results for a specific

region: a) to perform a small volume correction (SVC) and b) to perform a region of interest (ROI) analysis. For a SVC analysis a mask is used to define the area of interest and it is applied on already estimated whole-brain results. Significantly activated voxels are identified only within this small region and in this way the multiple-comparisons problem is less severe. The SVC analysis mask has to be independently identified so that the results will not be biased. In our case we defined the volume of interest anatomically, using the cytoarchitectonic probability map for left hemisphere BA44 provided in the SPM5 Anatomy toolbox (Eickhoff et al. 2005).

The second type of region specific analysis is the ROI analysis. For this analysis, the average signal of all the voxels included in the mask is computed and as a result, only one measurement is used per region, i.e. the average across all voxels. There are many ways to compute the average such as simple averaging, weighted averaging or principal component analysis (PCA). The only difference between using the mean and the first eigenvector from the PCA is that the mean is sensitive to outliers and the presence of noise. In our case we did not have any reason to think that there would be such outliers in the voxels included in the mask and a comparison between the mean and the first eigenvector did not reveal any differences either. Therefore in most cases, unless otherwise mentioned, we used the mean to average across voxels within the ROI. After the average is computed, the parameter estimates are calculated and submitted to statistical analyses. This approach followed the implementation of random effects analyses in the Marsbar SPM toolbox (Brett et al. 2002).

So far, we have described the most common methods used for the univariate analysis of the data. As previously mentioned, the exact details of the analyses performed in each study can be different depending on the study. Accordingly, we have included separate methods subsections in each of the chapters that describe the experiments performed. These subsections describe the precise details of the analyses applied to that experiment.

3.2.3.2 Psychophysiological Interactions

Another approach that we used is based on the principles of functional connectivity (Friston 1994; Horwitz et al. 1999). In short, functional connectivity refers to the correlation in the BOLD fMRI signal of two or more areas. Even though it is true that correlation is not causation and functional connectivity methods cannot provide information about the type of connection between the correlated regions, it can provide insights about the functional networks that are formed during a task and the changes in connectivity as a function of the changes in the experimental conditions. We were particularly interested in the latter, and used the SPM5 implementation of functional connectivity (Friston et al. 1997; Gitelman et al. 2003). In SPM5 they refer to this approach as a psychophysiological interaction (PPI). The idea is that responses in one cortical area can be explained in terms of an interaction between the influence of one area and some experimental (task-related in our case) parameter (Friston et al. 1997). This approach is different from a simple correlation in many ways, including the fact that the contribution is directional, i.e. from one seed region to a target region, and that in order to estimate this contribution the general linear model approach is used. Therefore, the presence of a significant interaction means that the contribution of one area to another changes significantly with the experimental factor.

As previously mentioned to calculate the PPI we use a linear regression. Simply put, the activity of a target region is regressed on the activity of a seed region and the slope of the regression reflects the influence that the seed region exerts over the target region (Friston et al. 1997). However, this measure is taken to be context-dependent and under different experimental conditions the slope changes. This change in the slope is what is referred to as a psychophysiological interaction. The statistical model used to estimate the interaction on target region *i* is:

$$x_i = x_k \times g_p \cdot \beta_i + [x_k g_p G] \cdot \beta_G + e_i \qquad (2)$$

The term $x_k \times g_p \cdot \beta_i$ is the interaction between the physiological activity in seed region *k* and some experimental parameter g_p, with β determining the strength of the interaction. *G* is the part of the design matrix that contains uninteresting effects/confounds and the term $[x_k g_p G] \cdot \beta_G$ is basically used for adjusting the data, i.e. removing the main effects of the seed region and the experimental factor, as well as other confounds. E_i is the error-term (Friston et al. 1997).

Measuring this interaction is therefore quite simple and based on linear regression. However, an added complication when using PPI on fMRI data is the fact that the relationship between the measured signal x_k and the actual neuronal response xn_k is filtered by the HRF. In essence, x_k equals xn_k convolved by the HRF. However, it has been shown that the product of the experimental factor convolved by the HRF and the measured BOLD signal, does not equal the product of the neuronal signal and the experimental factor (the product) convolved by the HRF (Gitelman et al. 2003).

$$g_p \times HRF \cdot x_k \neq (g_p \cdot xn_k) \times HRF \quad (3)$$

Since we are only able to measure the BOLD-fMRI signal, in order to get a better estimate of the interaction between the two factors, the experimental and the physiological, we need to deconvolve the BOLD-fMRI signal and derive an estimate of the underlying neuronal activity. This estimate can then be used to calculate the interaction with the experimental factor. At a final step the interaction is convolved with HRF and is used for regression with the measured signal of the target region x_i.

This is a brief presentation of the methods used for the functional connectivity analysis. Once again, more details on the conditions that were tested and the regions used are presented in the methods sections of the relevant chapters.

Chapter 4: From Phonemes to Articulatory Codes: an FMRI Study of the Role of Broca’s Area in Speech.

In two recent models of the neurophysiology of language (Hickok and Poeppel 2004; Indefrey and Levelt 2004), Broca's area was associated with different functions. As discussed in chapter 1, where we reviewed current models on language production, Indefrey and Levelt hypothesized that Broca's area was engaged at the level of phonological processing. In particular, they proposed that it is associated with the process of syllabification, one of the necessary steps prior to the retrieval or generation of the articulatory codes. In contrast, in the model proposed by Hickok and Poeppel, Broca's area was assigned to the next step after phonological processing, i.e. phonetic encoding and the mechanism of retrieving or generating the articulatory codes. In the present study, we address this issue and identify the level of processing the LIFG is involved in, phonological or phonetic. We used event-related functional magnetic resonance imaging (fMRI) and manipulated the phonological properties of pseudowords in a way that separated the processes of phonological and phonetic encoding. This manipulation allowed us to identify the key areas involved in the two levels of encoding and to disambiguate the function of Broca's area with respect to these two levels. We found significant activation of a premotor network consisting of the dorsal precentral gyrus, the IFG bilaterally and the supplementary motor area for low vs. high sub-lexical frequency pseudowords. We discuss our findings with respect to the mechanisms of phonetic encoding and generating articulatory codes and provide evidence in support of a functional segregation of the posterior part of Broca’s area, the pars opercularis. We conclude that the LIFG could have a role in both phonetic and phonological encoding, with different subregions underlying the different processes.

4.1 Experimental Hypothesis

Neuroanatomically, the processes of generating lexical phonological representations have been associated with the middle and posterior superior temporal gyrus (Fiez et al. 1999; Indefrey and Levelt 2000; Hickok and Poeppel 2004) also known as Wernicke's area. In some theories (Zatorre et al. 1996; Poldrack et al. 1999; Burton et al. 2000), they have also been assigned to Broca's area and specifically to the posterior, opercular part of the LIFG, roughly corresponding to Brodmann area 44 (BA44). As we have already mentioned in chapter 1, BA44 is thought to be specifically involved in syllabification (Indefrey and Levelt 2000) and sub-lexical processes that require explicit segmentation, such as tasks where subjects perform phonological decisions like phoneme monitoring, phoneme discrimination, or phoneme sequencing (Zatorre et al. 1992; Demonet et al. 1996; Zatorre et al. 1996; Poldrack et al. 1999; Burton et al. 2000). In the model proposed by Indefrey and Levelt (2004), the LIFG is part of a network related to segmenting a retrieved phonological word, while the premotor cortex (BA 6) is responsible for compiling and storing the motor codes for the individual syllables. Hence, according to this view, the premotor cortex is identified as the location of the mental syllabary (Levelt and Wheeldon 1994), rather than the LIFG.

This view is in contrast to the theory developed by Hickok and Poeppel (2004; 2007), which we also presented in chapter 1. Based on the Hickok and Poeppel model, Broca's area is part of the sensory-motor integration interface and in this sense it is directly involved in the generation or retrieval of the articulatory codes. The proposed role of Broca's area (along with the ventral premotor cortex) is to hold a speech sound map, i.e. representations of phonemes or frequent syllables and their associated motor programs (Guenther et al. 2006). This view is in contrast to the view of Indefrey and Levelt. according to their model the role of Broca's area is to support syllabification and post-lexical phonological processing. In contrast, Hickok and Poeppel propose that the role of Broca's area is related to phonetic encoding and

the generation of the articulatory scores, since it serves as a store for articulatory representations.

For this study we investigated the role of Broca's area in the process of generating an articulatory motor plan. We specifically wanted to address whether the posterior part of Broca's area (pars opercularis) is involved in (1) phonological processes, such as syllabification, or (2) in directly retrieving/compiling the articulatory gestures. The two hypotheses make different predictions about the sensitivity of the region to sub-lexical frequency effects. If the posterior part of Broca's area is only involved in the process of syllabification, it should not show a significant effect for sub-lexical frequency manipulations (prediction 1). Sub-lexical frequency effects are related to the process of phonetic encoding and accessing the articulatory codes for a particular target (Cholin et al. 2006). Based on the Indefrey and Levelt model (2000; 2004), at the stage of syllabification/phonological encoding the information on the articulatory codes are not yet available. Therefore regions who only involved in phonological processing should not be modulated by differences in sub-lexical properties.

However, if the Indefrey and Levelt model is false and the Hickok and Poeppel theory is correct (2000; 2004) we would expect that Broca's area would be involved in syllable articulatory code production. If so, then we would also expect that there would be a significant difference between high and low sub-lexical frequency items in Broca's area. Based on the theory on the existence of a mental syllabary or speech sound map, frequently used syllables would be pre-compiled and stored in the area, while infrequent ones would need to be compiled on-line based on their segmental features (i.e. phonemes). We would therefore expect to see higher activation for low vs. high sub-lexical frequency syllables in cortical areas that are involved in compiling the articulatory scores (prediction 2).

To address these questions we used event-related fMRI to monitor the changes in blood oxygenation while subjects performed a delayed phonological word repetition task. During the delay period, the subjects were given specific instructions to

rehearse the target stimulus covertly. After the delay period, an auditory probe instructed them as to whether they should repeat the presented word overtly or covertly. These instructions ensured that the articulatory code would be fully generated during the delay period. The presented pseudowords were constructed so as to be different in both length and frequency of segments and syllables. This manipulation resulted in a 2 x 2 x 2 factorial design with factors length (four vs. two syllables), sub-lexical frequency (low vs. high sub-lexical frequency) and response type (overt vs. covert). During the construction of the stimuli, we also controlled for phonological neighbourhood density so that none of the pseudowords presented had any immediate phonological neighbours. This ensured that during the performance of the task we would not see any differences related to lexical effects but only related to the processes of generating articulatory codes. As previously mentioned, manipulating stimulus length and contrasting longer vs. shorter pseudowords would reveal the network underlying phonological and phonetic encoding. These processes are considered to be incremental (Levelt et al. 1999; Guenther et al. 2006) and longer targets have longer processing time.

In order to identify the areas specifically participating in compiling the articulatory codes for retrieved phonological codes, we also manipulated the sub-lexical frequency of the pseudoword components (for both syllables and phonemes). The regions who would show a significant activation for the contrast between low and high sub-lexical frequency stimuli would comprise the network underlying the generation of articulatory codes and participate in on-line articulatory code generation. The opposite contrast, high vs. low frequency pseudowords, would reveal the location of the mental syllabary. If Broca's area is involved in syllabification and phonological processing prior to the encoding of the articulatory scores, it would only show a strong effect of length, but not frequency. On the other hand, if Broca’s area is the site of the mental syllabary, we expected to see significant effects of both length and frequency manipulations.

4.2 Methods

4.2.1 Data Acquisition

Fifteen healthy, right-handed, monolingual native speakers of American English were chosen to participate in the study (8 males, 7 females) with mean age 26 years (range = 20-35). Two subjects (one male, one female) were excluded from analysis because of excessive head motion. The subjects laid in the fMRI scanner and were asked to perform a delayed, pseudoword repetition task. The presented pseudowords belonged to one of four experimental conditions: four-syllable low frequency, four-syllable high frequency, two-syllable low frequency and two-syllable high frequency.

Over the course of two experimental fMRI runs, subjects were presented auditorily with thirty-six items per condition (for a total of 144 items per participant). After a delay of 6 seconds, a probe (one of two versions of a bell sound) was heard instructing the subject to repeat the presented word either overtly or covertly (depending on the type of probe). During the delay period, the subjects were given specific instructions to rehearse the presented stimulus covertly. They also did not know prior to the presentation of the relevant probe whether they would be asked to respond overtly or covertly. Therefore, we expected that during the delay period they would fully retrieve the articulatory scores for the presented word. Each trial lasted 8 seconds (see for a diagram of the structure of the experimental trials Figure 9-A).

Stimulus presentation was in a pseudorandom, fast event-related fashion whereby the order of occurrence for the conditions was controlled by a combination of three shifted versions of a binary, 63-bin m-sequence (one shifted by 9 bins and the other by 18 bins with respect to the first one; an example is presented in Figure 9-B). This

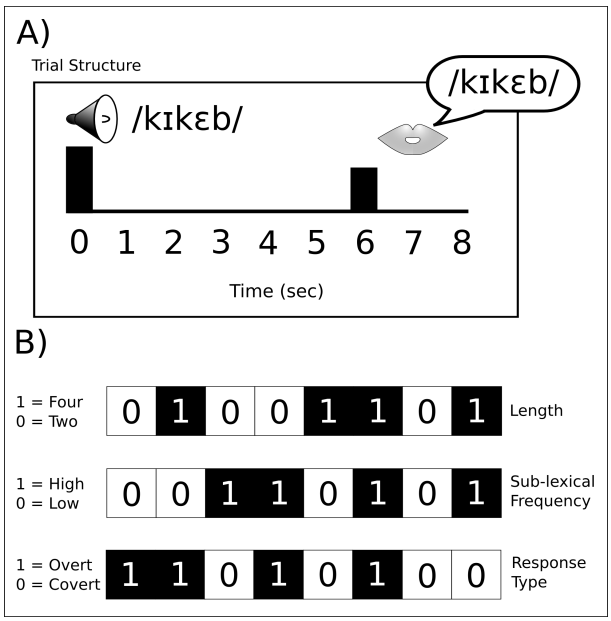

Figure 9: During the experiment, subjects were asked to listen to pseudowords and to repeat them either overtly or covertly after a 6s delay. The structure of each trial is shown in (A). The stimulus is presented auditorily at 0s and subjects then wait for the response probe. During the delay period, they are instructed to covertly rehearse the stimulus and are not aware of the type of response (overt or covert) before they hear the probe. The type of stimulus that will be presented in each trial is determined pseudo-randomly by a combination of 3 m-sequences. In (B) we present an example of 3 binary sequences that resemble those used in the experiment. Each sequence is associated with an experimental factor. In the example provided, the top sequence controls the length of the stimulus (1 for four syllables, 0 for two syllables), the middle sequence controls sub-lexical frequency (1 for high, 0 for low) and the bottom sequence controls response type (1 for overt, 0 for covert). For example, the combination 0 1 0 would result in the presentation of a two-syllable, high-frequency pseudoword and the covert response probe.

ensured that the experimental conditions would be orthogonal to one another and counterbalanced (Buracas and Boynton 2002; Kellman et al. 2003). The binary m-sequence was padded in the beginning with 9 more trials (for a total of 72 trials), which were not analysed for the purposes of this study.

Prior to the onset of the experiment, all subjects performed a 15 minute practice session outside the scanner to allow them to become familiar with the structure of the task and its demands. A quality check run was also performed prior to the onset of the experimental runs. During this run, the volume of the headset was adjusted based on the subject's feedback and the images were checked to make sure that they were free from artifacts.

As mentioned in chapter 3, imaging was performed on a 3T MRI system using single-shot rate-2 SENSE EPI for the acquisition of the functional scans (de Zwart et al. 2002). The scanning parameters used were TE=31ms, flip angle of 90 degrees, TR=2s and acquisition bandwidth 250 kHz. A total of 32 axial slices were acquired interleaved with slice thickness = 2mm (gap = 0.3mm) and an in-plane resolution of 2.3x2.3mm^2 (96x72 matrix, 22.4x16.8cm^2 FOV). Four volumes were acquired during each trial. Because of the smaller-than-usual thickness of the slices we could not image the whole brain and acquired images in a slightly oblique position, covering an area from below the STS to the top of the head. At the end of the scanning session, high-resolution spin-echo T1 anatomical images were acquired at the same location as the functional EPI scans. The scanning parameters for the anatomical image were: TR=700ms, TE=13ms, 256x192 data matrix with a 22.4x16.8cm^2 FOV and 2mm slice thickness (with 0.3mm gap), resulting in 0.86x0.86mm^2 voxel size.

4.2.2 Data Preprocessing and Analysis

All analyses and image preprocessing were carried out using the SPM5 software package and associated toolboxes (http://www.fil.ion.ucl.ac.uk/spm/software/spm5). Details on the the preprocessing and motion correction protocol followed are reported in section 3.2.1. The analysis of the realignment parameters with respect to scan-to-scan motion (a three factor ANOVA with within-subject factors response type, stimulus length and sub-lexical frequency) revealed a significant main effect of response type in all directions ($F(1, 12) > 26$, $p < 0.004$ for all directions). In agreement with other studies (Barch et al. 1999; Shuster and Lemieux 2005), the incremental movement was overall quite small and greater for overt response trials (mean ±std displacement was 0.039mm ±0.014 for translations and 0.034º ±0.012 for rotations) than covert response ones (mean ±std was 0.02mm ±0.008 for translations and 0.017º ±0.006 for rotations).

Additional significant effects were present for length in the roll rotation and for both

the main effect ($F(1, 12) = 5.9$, $p < 0.04$) and the interaction between length and response type ($F(1, 12) = 19$, $p < 0.001$). Four-syllable pseudowords (mean roll displacement was 0.038º ±0.02) produced greater movement than two syllable pseudowords (mean was 0.034º ±0.02) and especially during overt responses. Finally, in the y direction there was a significant main effect of sub-lexical frequency ($F(1,12) = 6.3$, $p < 0.03$) and interaction between sub-lexical frequency and response type ($F(1,12) = 10.8$, $p < 0.01$). Low frequency items caused more movement (mean 0.021mm ±0.013) than high frequency items (0.019mm ±0.010), especially during overt response trials.

As reported in chapter 3, the realignment parameters were included in the design matrix as effects of no interest and an additional regressor was added for images that showed changes in the global signal greater than 1.5% of the mean followed by a greater than 0.5mm incremental movement (Mazaika et al. 2007). Finally, we inspected the movement parameters (both incremental and absolute motion) and excluded from the analysis two subjects that showed incremental movement greater than the criteria we set in chapter 3, i.e. motion greater than 1mm or 1º. All subjects met the absolute motion inclusion criteria.

Statistical analysis of the factorial event-related experiment was performed in SPM5 using the FIR approach as described in chapter 3. In summary, a 3-way, within-subject ANOVA was performed with factors length (four- vs. two-syllable pseudowords), sub-lexical frequency (low vs. high) and response type (overt vs. covert). The ANOVA was implemented in two levels as described in chapter 3. T-contrast images were produced with the use of two HRFs as the contrast vectors to model the presentation and delay period and the response period respectively. Statistical parametric maps (SPM) were thresholded at $p < 0.001$ uncorrected at the voxel level and $p < 0.05$ corrected for family-wise error (FWE) at the cluster level (Hayasaka and Nichols 2003). For this study, significant clusters had on average more than 85 voxels.

In order to analyse the contrast estimates for the LIFG, we used the cytoarchitectonic probability map for left hemisphere BA44 (Eickhoff et al. 2005). For each of the main effects of interest (length, frequency and response type), we identified the voxels within the activated clusters that were part of BA44. We then extracted the average beta weights (over cluster voxels) for each of the four conditions of interest in the design (4 syllable low frequency, 4 syllable high frequency, 2 syllable low frequency and 2 syllable high frequency) and for all subjects. A single value corresponding to the weighted sum of the estimates across the FIR (weighted by the HRF) was then extracted for each of the four conditions and subjects and used in multiple 2-sided t-tests testing for effects of frequency, length or the difference between the two conditions within each region. This approach followed the implementation of random effects analyses in the Marsbar SPM toolbox (Brett et al. 2002). Significance was determined using a threshold of $p < 0.05$. Where appropriate (more than one ROI) the p-values were adjusted to correct for multiple comparisons (Bonferroni correction).

Finally, we examined the connectivity changes as a function of sub-lexical frequency using a psychophysiological interaction analysis (PPI) as is implemented in SPM5 (Friston et al. 1997; Gitelman et al. 2003). This analysis revealed differences in connectivity between cortical regions during the processing of e.g. low vs. high frequency pseudowords. Because we did not have an a priori hypothesis about which cortical connections would change as a function of sub-lexical frequency we performed a mass-univariate connectivity analyses and examined the correlation between specified seed regions and the rest of the cortex. To identify the regions that would be used as “seeds” and for which the connectivity with the rest of the cortex would be calculated, we used the results from the subtraction analysis. For each of the clusters of interest and for each subject we identified the activation peaks and extracted the BOLD signal time-series data averaged over a sphere with 5mm radius around the activation peak. Subjects who did not show significant activation in the specified regions above a threshold of $p < 0.1$ uncorrected were excluded from the analysis. After the (PPI) vectors representing the interaction between the

psychological and the physiological factors were estimated (for details of the estimation process see section 3.2.2.2), they entered a regression analysis. The regression slope determined the direction of the connectivity between the contrasting conditions, e.g. for the contrast low vs. high frequency a positive slope means that the correlation between the seed and the target region is more positive during the processing of low frequency syllables than high. Significance was determined at $p < 0.001$ uncorrected at the voxel level and $p < 0.05$ FWE corrected at the cluster level (Hayasaka and Nichols 2003). For this analysis, significant clusters had on average more than 45 voxels.

4.3 Results

4.3.1 Behavioural Results

To test for effects of length or frequency on subject performance we measured subject response accuracy. Based on previous results, we expected to find a decrease in response accuracy for low frequency pseudowords, but we did not expect to find an effect of length. We performed a 2-way ANOVA with length and sub-lexical frequency as within-subject factors. As expected we found that there was a significant main effect only for the frequency condition ($F(1, 12) = 14.62$, $p < 0.003$). No other main effects or interactions were significant. Mean (±std) accuracy rates were 64.5% (±15) for low frequency pseudowords and 75% (±13) for high frequency pseudowords. The relatively low accuracy scores were expected, considering the nature of the task (pseudoword repetition) and the noisy environment. However, all subject performance accuracy was within three standard deviations of the group mean (70%, std = 13).

Finally, to ensure that there was a significant difference in sub-lexical frequency between the responses, we calculated the phoneme and biphone phonotactic

probability (PP) of the subject overt responses and performed a two-sided t-test to compare high vs. low frequency responses. For both biphone and phoneme measurements, the differences were significant ($t(12) = 14.66$, $p < 0.001$ for biphones and $t(12) = 15.74$, $p < 0.001$ for phonemes). Mean (±se) PP scores for high frequency responses was 0.0193 (±0.0009) for biphones and 0.3656 (±0.0145) for phonemes. Low frequency PP scores were 0.0025 (±0.0006) for biphones and 0.1187 (±0.0091) for phonemes. From the above results, we can conclude that the subjects perceived the differences between low and high frequency targets and performed the task according to the instructions.

4.3.2 FMRI Results

4.3.2.1 Phonological Encoding

To map the areas involved in phonological encoding we compared the activation levels invoked for processing four- vs. two-syllable pseudowords (over both low and high frequency syllables). A significant main effect of length (four- greater than two-syllable stimuli) was observed in a large perisylvian network extending bilaterally across the superior temporal gyrus (STG), the precentral gyrus (PrCG) and the pre-supplementary motor area (pre-SMA), as well as the left inferior frontal gyrus (LIFG) (cf. Figure 10-A for whole brain results and Figure 10-C for significantly activated voxels within the LIFG). The largest activations were observed in the left hemisphere (L) for a cluster that covered both the PrCG and STG. In particular for the STG, the cluster covered a large portion of the middle and posterior STG including the upper banks of the superior temporal sulcus (STS) and an area in the junction between the parietal and temporal lobe also referred to as Sylvian parieto-temporal area (Spt) (cf. Table 2 for the coordinates of the significantly activated areas). The LSTG has been previously implicated in phonological processing (Indefrey and Levelt 2000; 2004; Graves et al. 2007), while the LPrCG is a known

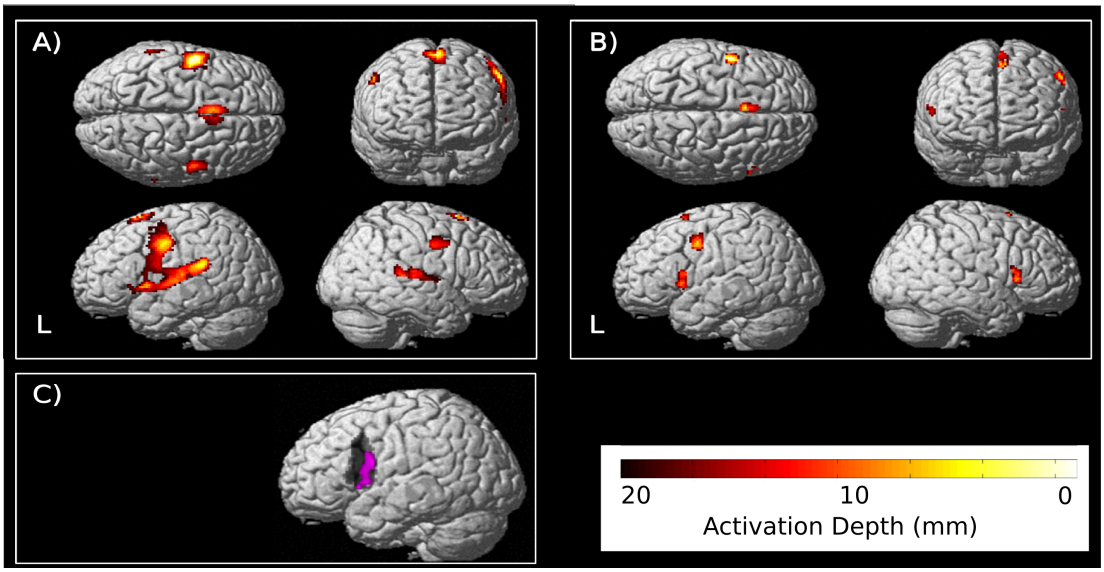

Figure 10: Surface renderings of significant activations in the whole-brain random effects analysis for length (A) and sub-lexical frequency (B). In (A) the contrast four vs. two syllables yielded significantly higher activation in perisylvian and premotor regions including the LIFG. In (B) premotor areas including the dorsal precentral gyrus and the IFG bilaterally showed significantly higher activation for low vs. high frequency pseudowords. In (C) we show the main effect of length within left BA44 (significantly activated voxels appear in magenta) using a small volume correction approach (SVC). BA44 (shaded area) was defined using a cytoarchitectonic probability map of the area (Eickhoff et al. 2005). Maps are thresholded voxel-wise at $p < 0.001$ uncorrected and cluster-wise at $p < 0.05$ FWE corrected. Colour grading in (A) and (B) reflects depth of the supra-threshold voxels, with brighter voxels on the surface. The maximum depth of the projected voxels is 20mm. LIFG, left inferior frontal gyrus; L, sagittal view of the left hemisphere.

premotor area and as such it has been associated with phonetic encoding. A similar effect could also be observed for the LIFG. The activated area was located on pars opercularis and ran along the inferior frontal sulcus (IFS). In accordance to our hypothesis, we expected the result that both phonological and phonetic encoding processes would show an effect of length. What distinguishes the two processes is their sensitivity to sub-lexical frequency. If a region is involved in phonological processing, we would not expect it to show significant sub-lexical frequency effects (prediction 1). On the other hand, if it is involved in phonetic encoding, we would expect it to show significant effects for both conditions, length and sub-lexical

frequency (prediction 2).

Table 2: Brain Regions Modulated by Length and Frequency

Contrast	Region	Coordinates			T	Size
		x	y	z		
Four > Two Syllables	Left precentral gyrus	-56	-4	44	7.87	2097
	* Left superior temporal gyrus	-60	-12	4	6.76	
	* Left sylvian parieto-temporal junction	-56	-38	20	5.82	
	* Left inferior frontal gyrus	-60	4	20	4.63	
	Left pre-supplementary motor area	-4	10	68	7.21	388
	Right superior temporal gyrus	50	-22	8	5.45	393
	* Right sylvian parieto-temporal junction	64	-32	10	5.24	
	Right precentral gyrus	50	-4	40	5.30	176
Low > High Frequency	Left precentral gyrus	-52	2	40	4.77	138
	Left pre-supplementary motor area	-4	14	58	4.51	122
	Left inferior frontal gyrus	-54	12	12	4.01	119
	Right inferior frontal gyrus	50	18	4	4.23	97

*Note: Regions significantly activated in the random-effects group analysis ($t(144) > 3.1$, $p < 0.05$ FWE corrected for cluster-size). Displayed are the contrasts, the coordinates for the voxels of greatest activity within the activated clusters in MNI stereotactic space, a description of the region, the T value and the size of the activated cluster (in number of voxels). In the case of very large clusters, multiple peak voxels are reported. These are prefixed with a * and they are clustered together with the last non-prefixed entry in the table.*

4.3.2.2 Phonetic Encoding

Comparing pseudowords with low vs. high phonotactic probability syllables and segments revealed regions that showed an effect for sub-lexical frequency. Based on our hypothesis, areas that showed a frequency effect reflect the process of phonetic encoding, i.e. articulatory code generation (Indefrey and Levelt 2000). Four regions showed significant main effects of frequency: the left hemisphere dorsal PrCG, the left hemisphere pre-SMA and the IFG bilaterally (cf. Table 2 for a detailed list of the activated regions and Figure 10-B for a map of the significantly activated areas). Activity in the LSTG did not reach significance ($p < 0.3$ FWE corrected voxel-wise , $p < 0.2$ FWE corrected cluster-size), which is consistent with a role of this area in phonological rather than phonetic processing (prediction 1).

We also tested for the opposite contrast, high vs. low frequency pseudowords, in order to see whether the areas associated with retrieving high-frequency, pre-compiled syllables from the mental syllabary, are different from the ones associated with on-line generation of articulatory scores. No areas showed higher activation for high vs. low frequency syllables. There were also no significant interaction effects between length and sub-lexical frequency.

In addition to the subtraction analysis, we also performed a PPI analysis to identify the connectivity changes as a function of sub-lexical frequency. We were particularly interested in the changes in connectivity between regions such as left hemisphere IFG and PrCG. We were interested in seeing whether the differences in the processing of high vs. low sub-lexical frequency pseudowords would also be associated with differences in connectivity. We defined the seed regions using the activation peaks reported above for the contrast low vs. high sub-lexical frequency and looked at the differences in the PPI between the time-course of the seed regions and the rest of the imaged cortex. We observed a significant decrease in the connectivity only for the left hemisphere PrCG seed, with high frequency pseudowords evoking stronger connectivity than low frequency pseudowords. More

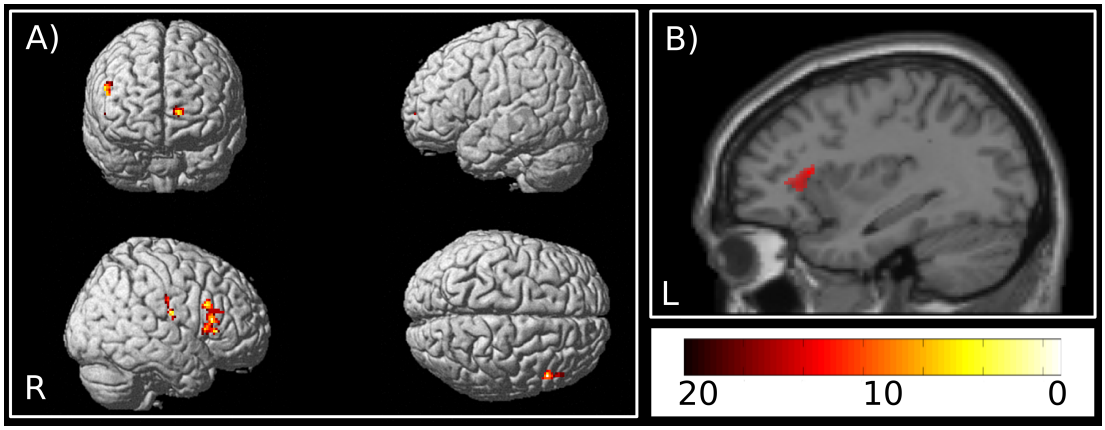

Figure 11: Statistical maps of significant changes in connectivity for high vs. low sub-lexical frequency. The seed region used was in the left PrCG (MNI -52 2 40). In (A) is a surface rendering of the results. Because one of the regions, the FOI, was too deep and could not be rendered clearly on the surface templates, in (B) we present a sagittal view of the region at x = -34. Maps are thresholded voxel-wise at p < 0.001 uncorrected and cluster-wise at p < 0.05 FWE corrected. Colour grading in (A) reflects depth, with brighter voxels on the surface. The colour-bar on the bottom right is in mm and shows the depth of the supra-threshold voxels. The maximum depth of the projected voxels is 20mm. L, sagittal view of the left hemisphere; R, sagittal view of the right hemisphere; FOI, frontal operculum and anterior insula junction; PrCG, precentral gyrus.

specifically, the left PrCG (MNI -52 2 40) decreased its connectivity to a region in the junction between the left frontal operculum and the anterior insula (FOI; MNI -34 24 12; t(10) = 9.45, 78 voxels, p < 0.003 FWE corrected cluster-wise), the right IFG (MNI 50 18 26, t(10) = 8.54, 221 voxels, p < 0.001), the medial part of the left superior frontal gyrus (mSFG; MNI -12 54 6; t(10) = 7.10, 46 voxels, p < 0.04 FWE-corrected cluster-wise) and finally, the right rolandic operculum (RO; MNI 42 -10 18; t(10) = 5.64, 92 voxels, p < 0.002 FWE-corrected cluster-wise; see Figure 11).

To further understand the nature of the differences observed in the PPI (if one slope is zero or negative) we also examined the regression slopes for each condition. For all target regions, the regression slope for low sub-lexical frequency (mean ±se was 0.49 ±0.03), which was lower than for high sub-frequency pseudowords (0.68

±0.06), but it was neither zero nor negative. Finally, we also inspected the PPI maps for significant increases in connectivity for low vs. high sub-lexical frequency. For the seed regions used, we did not find any significant results, i.e. stronger connectivity for low frequency pseudowords as compared to high frequency ones.

4.3.2.3 Left Inferior Frontal Gyrus

To further test our hypothesis about the involvement of Broca's area in phonetic processing, we performed an ROI analysis. A region corresponding to the LIFG was independently identified using the contrast overt vs. covert repetition (centre of mass x = -55, y = 9, z = 13, size = 138 voxels). In a random effects two-way ANOVA with factors length (four vs. two syllables) and sub-lexical frequency (low vs. high) the LIFG showed a main effect for both factors ($t(12) = 3.5$, $p < 0.003$ and $t(12) = 2.2$, $p < 0.03$ for length and frequency respectively).

Because the LIFG showed effects for both length and frequency we further investigated whether there were any signs of functional segregation within the IFG and in particular the pars opercularis, as had been observed in other studies (Molnar-Szakacs et al. 2005). For the main effect of length and sub-lexical frequency, we observed two clusters within the LIFG, which were only partly overlapping (9 voxels out of 82 and 79 respectively for the two clusters; Figure 12). The distance between their centre of mass was 9 mm, i.e. a factor of 1.5 greater than the smoothing kernel (6mm), with the cluster showing a greater effect of length following the anterior banks of the precentral sulcus and extending more lateral, posterior and dorsal to the cluster showing a greater effect of frequency. We will refer to the cluster identified during the length condition as dPOp (dorsal pars opercularis) and the cluster identified for the frequency condition as vPOp (ventral pars opercularis), because of their anatomical differences and in agreement with previous evidence.

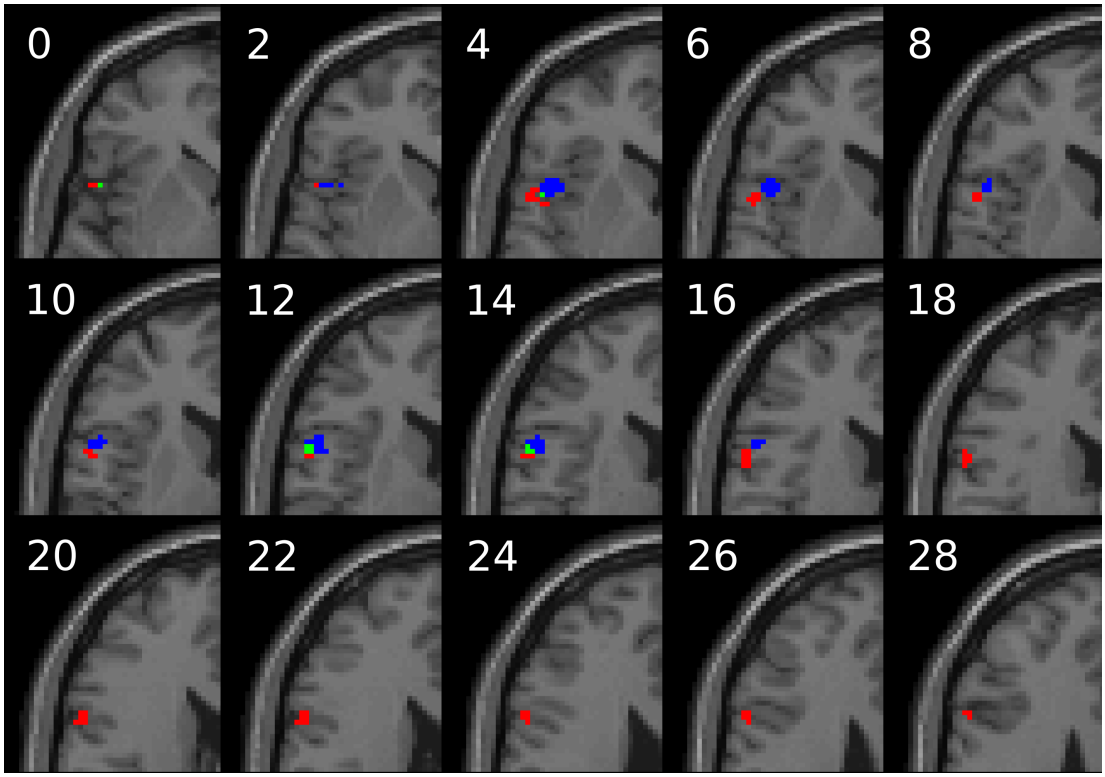

Figure 12: Significant activations within left hemisphere BA44 as defined by a cytoarchitectonic probability map of the area (Eickhoff et al. 2005). Shown in red are significantly activated voxels for four vs. two syllables. This cluster extends from z=-2 (slice not shown) to z=28. The highest activation is located dorsally, at [-60 4 20]. Shown in blue are significantly activated voxels for low vs. high sub-lexical frequency. The highest activation is located at [-54 12 12]. Finally, shown in green are voxels that are overlapping for both conditions (size of overlap = 9 voxels). Activations are thresholded at p < 0.05 FWE corrected voxel-wise. Coordinates are in MNI space.

Both the dPOp and the vPOp exhibited effects of frequency and length, though the frequency effect for dPOp was just slightly below threshold (dPOp frequency: $t(12) = 2.5$, $p < 0.06$; vPOp length: $t(12) = 3.2$, $p < 0.02$ corrected for two ROI). This difference already suggests that there might be a functional segregation within the pars opercularis of the LIFG. To further examine whether there is a functional difference in the activation between the two clusters, we examined the region (dPOp vs. vPOp) by experimental condition (length vs. frequency) interaction (Friederici 2006). We performed a 2-sided paired t-test on the region-specific differences

between the length and frequency conditions and found a significant region-by-condition interaction ($t(12) = 3.1$, $p < 0.01$), indicating that there is a robust difference between the two clusters in terms of their response to length and sub-lexical frequency effects. DPOp shows greater activation for length rather than sub-lexical frequency (mean ±se length over frequency difference is 0.093 ±0.051), while in vPOp there is almost no difference between the levels of activation for the two conditions (mean ±se length over frequency difference is 0.002 ±0.026).

4.4 Discussion

In this study we were able to delineate the cortical areas involved in the phonemic to articulatory translation that is necessary for the generation of articulatory codes. By directly contrasting targets with varying length, we manipulated the load on the system of articulatory-motor production and were able to identify a number of key regions underlying articulation and the overall process of transforming phonological word forms to articulatory codes. In summary, these regions included bilateral (although strongly left lateralized) mid and posterior superior temporal and frontal regions, the premotor cortex and the pre-supplementary motor area. These results are in agreement with current models on word production that describe a left-lateralized, perisylvian network (Indefrey and Levelt 2000; 2004; Hickok and Poeppel 2004; 2007).

To further identify the roles of the different components of the network, and in particular to resolve the conflict on the role of the left inferior frontal gyrus, we probed the network by manipulating sub-lexical frequency. Our hypothesis was that only regions that are directly involved in phonemic-to-articulatory translation would show an effect for frequency manipulation. Targets with high-frequency components (whether we consider syllables or phonemes as the structural unit[8]) are processed

8) For evidence in support of syllables see Levelt and colleagues, (Levelt et al. 1999; Cholin et al. 2006)

faster than the ones with less frequent components (Vitevitch and Luce 1998; 2005). Such evidence suggest that targets with components of different sub-lexical frequency (high vs. low) are processed differently (Guenther et al. 2006). High-frequency clusters are pre-compiled and their articulatory codes need to be retrieved, while low-frequency clusters need to be compiled on-line on a segment-to-segment basis (Guenther et al. 2006).

In our experiment we identified four regions that showed an effect related to sub-lexical frequency (higher activation for low vs. high frequency): the left hemisphere pre-SMA, the left hemisphere PrCG and the IFG bilaterally. From previous studies on motor planning and production, it is known that the SMA have a role in motor planning and the preparation of movements. Even though its function is not strictly associated with linguistic processes, it is also part of linguistic motor planning (Riecker et al. 2005). It has been shown that the rostral part of the SMA (pre-SMA) contains cells that code for an entire sequence to be produced, which in our case would correspond to a syllable or a sequence of syllables. In a recent fMRI study, the pre-SMA was shown to be sensitive to sequence complexity effects both within and beyond the syllable boundaries (Bohland and Guenther 2006). The present findings are in agreement with the current theories about the function of the pre-SMA and the observed frequency effect could simply represent the higher system processing load that is associated with processing new and unfamiliar motor plans (low sub-lexical frequency pseudowords) compared to familiar, more rehearsed and possibly pre-compiled ones (high sub-lexical frequency pseudowords).

The significant activation difference for low vs. high sub-lexical frequency pseudowords in the left precentral gyrus is also in agreement with current models on word production (Hickok and Poeppel 2004; Indefrey and Levelt 2004; Guenther et al. 2006). It is worth highlighting that only a small area in the dorsal PrCG was significantly active and that this area has been previously involved in studies examining sensory-motor mapping (Hickok and Poeppel 2004). Hickok and Poeppel talked about a “dorsal stream” in speech processing, which is involved in mapping

sound onto articulatory-based representations. The regions that are part of this stream include a posterior inferior frontal area (including Broca's area), a more dorsal premotor site and an area in the posterior parietal lobe, deep within the Sylvian fissure and at the boundary between the parietal and temporal lobes, also known as the Sylvian parieto-temporal junction, or Spt (Hickok et al. 2003). Area Spt, which lies within the boundaries of the planum temporale (PT), an area traditionally associated with acoustic and phonological processing, is thought be involved in speech production and to be the interface for the sound-to-gesture transformation. In our study, this area showed significant effects for target length, but not sub-lexical frequency. In our task we cannot distinguish between the processes of generating a phonological representation during perception of a presented target and generating a phonological representation for articulatory rehearsal, which could be either separate (Indefrey and Levelt 2004) or common, as suggested by the motor theory of speech perception (Liberman and Mattingly 1985). It is therefore not possible for us to say whether the activation in Spt is related to stimulus presentation, motor planning or both. However, the absence of significant frequency effects from this region highlights the fact that if this region is involved in sensory-motor mapping, then its role is likely to be related to sub-lexical phonological processes, such as syllabification and segmentation in preparation for generating the articulatory codes. This claim would be in agreement with older claims made by Indefrey, whereby a portion of the superior temporal lobe was considered as a possible candidate region for syllabification (Indefrey and Levelt 2000). The other candidate was the left inferior frontal gyrus.

In our study we also found significant activation in the LIFG. In particular, the pars opercularis, which roughly corresponds to BA44 (Amunts et al. 1999),showed consistent effects for both length and sub-lexical frequency (four vs. two syllables and low vs. high frequency, respectively). Furthermore, we found that there is a functional segregation within the pars opercularis. The more dorsal part of the area (dPOp) is modulated by differences in stimulus length, while the ventral part (vPOp) is modulated by differences in both length and sub-lexical frequency. The idea that

Broca's area is functionally segregated into its three anatomical parts (pars opercularis, triangularis and orbitalis) is well known and well founded (Bokde et al. 2001; Devlin et al. 2003; Heim et al. 2003a). Recently, however, there have also been claims about a functional segregation within the pars opercularis (Chein et al. 2002; Molnar-Szakacs et al. 2005). In a meta-analysis of imaging studies on imitation and action observation, Molnar-Szackacs et al. identified two distinct foci within the pars opercularis, a dorsal and a ventral one, that serve different functions. DPOp shows mirror neuron properties and is significantly active during both action observation and imitation, while vPOp shows only motor properties and is only active during imitation.

In agreement with this segregation, we also found two functionally segregated clusters in the pars opercularis with one extending more dorsally than the other. The more dorsal cluster is located closer to the inferior frontal sulcus and the premotor cortex and shows greater activation for length manipulation when compared to the vPOp, which is also significantly activated for low vs. high frequency stimuli. In our study, the dPOp is part of a wider area of activation in the left hemisphere PrCG. Therefore, based on its relation to premotor areas, as well as the fact that it is only active for the length condition, we conclude that the dPOp is involved in syllabification as has been proposed by Indefrey and Levelt (2000; 2004). This role is in agreement with other proposed roles such as sequencing discrete units (Gelfand and Bookheimer 2003) or sub-lexical processing requiring explicit segmentation (Zatorre et al. 1996; Burton et al. 2000; Chein et al. 2002).

The vPOp on the other hand shows a significant effect of both length and frequency, a finding that is in agreement with a role as the speech sound map or mental syllabary that has been proposed by Guenther et al. (2006). These results are partially in agreement with the claims made by Molnar-Szakacs and colleagues, whereby the vPOp is not a premotor region, in the sense that it is not directly involved in motor planning, but that it holds some form of representation of the motor plans that is communicated to the posterior part of the superior temporal sulcus (Molnar-Szakacs

et al. 2005). The exact hypothesis proposed by the authors suggests that the vPOp produces an efferent copy of the target motor plans that is sent to the STS during imitation allowing the prediction of the sensory consequences of planned imitative actions. However, the creation of a "copy" suggests that there is a target somewhere that shares the same characteristics as the copy. In our case, this would mean that the vPOp is not the location of the speech sound map as has been proposed, but that it holds a copy of the articulatory codes. The codes themselves are generated elsewhere. The only other possible candidate in our case would be the dorsal premotor cortex, which also showed a significant effect of sub-lexical frequency. Based on our results we cannot exclude either possibility.

Research into the functional segregation of the pars opercularis is still in its preliminary stages. In addition, the anatomy of the LIFG is very variable across subjects (Amunts et al. 1999), which makes it difficult to draw any precise conclusions about the exact anatomical borders of the hypothesized segregation of the pars opercularis based on group-averaged results. As imaging methods improve with high-field strength scanners and improvements in receive coil arrays, it is expected that the spatial resolution in fMRI will further improve to allow for more fine-grained differences to appear. For the purposes of this study, we have defined the two areas in gross anatomical terms such as ventral and dorsal based on the location of the activation peak within the clusters, which represents the group tendency. Future research using higher spatial resolution would be needed to further verify and specify the exact anatomical features of this functional segregation.

Regarding the subtraction analysis, we also note that we did not find any significantly activated regions for the inverse contrast high vs. low sub-lexical frequency. Based on our hypothesis, we would expect that a significant activation for this contrast would reveal the location of the mental syllabary versus the network underlying articulatory code generation. However, based on the computational model proposed by Guenther et al. (2006), the speech sound map does not just contain pre-compiled frequent syllables, but also motor representation for phonemes. The speech

sound map is therefore involved in both processes, though the on-line compilation of articulatory codes would be computationally more demanding than the retrieval of pre-compiled gestural scores. Therefore, it is not surprising that we do not see effects for high vs. low frequency stimuli, since it would be the same network that is underlying the process.

More information about the differences in the processing of high vs. low sub-lexical frequency stimuli is provided by the connectivity analysis. Even though based on the subtraction analysis, there were no regions that were significantly more active for high vs. low frequency stimuli, the PPI analysis revealed that the correlation of the PrCG with other cortical regions was significantly more positive for high rather than low frequency stimuli. In addition, there were no significant PPI results involving the LIFG. These results are at first surprising, since in a previous connectivity study (Bokde et al. 2001) it had been shown that the connectivity between the LIFG and the STG increases as a function of lexicality, with pseudowords showing the stronger connectivity when compared to real words. Their results were interpreted as an increase in effort to retrieve a lexical representation, which is non-existent for pseudowords, though the presence of a phonological neighborhood can still produce many candidates that need to be validated. However, in our case we are comparing within pseudowords and we do not expect that the results would show differences related to phonological neighborhood effects, since we have controlled for that. On the contrary what we think that the results reflect is a difference in the generation of the phonetic code. High sub-lexical frequency syllables are stored in the mental syllabary, while low frequency ones need to be compiled online.

We did not have a specific hypothesis about which regions would be affected by the differences in processing between the two conditions and used seed regions that showed significant differences in their activation during the subtraction analysis. We were then interested to see which other regions, if any, modulate this difference in activation. For a seed region in the LIFG, the connectivity analysis revealed no differences between processing low and high frequency stimuli. This means that the

regions interacting with the LIFG remain the same during both conditions. The only differences that we could observe are with respect to a seed region in the left precentral gyrus (LprCG). During the processing of high frequency pseudowords, this region increased its connectivity to regions in the prefrontal cortex (medial superior frontal gyrus; mSFG), the junction of the frontal operculum and anterior insula, the right rolandic operculum and the RIFG.

The stronger connectivity of the premotor area with a prefrontal region (mSFG) when processing high frequency pseudowords, possibly reflects the process of retrieving the precompiled articulatory scores from the mental syllabary. The junction of the anterior insula and frontal operculum has been previously shown to be involved in speech production and in particular to be sensitive to syllable complexity, both within and between syllables (Bohland and Guenther 2006). Based on these results, the authors suggested that this part of the cortex is involved either in "integrating affective and linguistic prosody" in the speech motor plan or it could also be a portion of the speech sound map. In our case, we did not expect any affective or prosodic differences between the two stimulus types.

However, the presence of a strong RIFG activation both in the subtraction analysis (higher for low frequency pseudowords) and in the PPI analysis (stronger connectivity with the PrCG for high frequency pseudowords), suggests that the two categories of stimuli might also be processed differently in terms of prosody. This difference cannot be perceived as related to the location of the stress, since there was no consistent difference in the stress pattern between the two categories of the presented stimuli. A possible explanation for the discord between the subtraction and PPI results is that intonation is easier to be retrieved and processed for high sub-lexical frequency syllables, which are more familiar to the system. In the case of low frequency components, intonational patterns are more unfamiliar and the increase in BOLD signal observed in the subtraction analysis could represent the difficulty in processing the intonational patterns of the low frequency pseudowords, in which case they may contribute less directly to the generation of the articulatory scores.

To conclude, in this fMRI study we investigated the processes of phonological-to-articulatory translation and the role of the left inferior frontal gyrus. Based on our findings, we conclude that the LIFG, BA44 in particular, is functionally segregated into two subregions, following a dorsal-ventral gradient. The dorsal part is involved in phonological segmentation, while the ventral part is involved in the translation between phonemic and articulatory representations. This finding is in agreement with recent observations on the functional segregation of the pars opercularis and further clarifies the role of the LIFG in language production.

Chapter 5: Phonetic Encoding vs. Working Memory: Is Broca's Area Necessary for Phonetic Encoding?

In the previous fMRI experiment studying the role of the left inferior frontal gyrus (LIFG), it was found that the LIFG is functionally segregated and its ventral part (vPOp) is sensitive to sub-lexical frequency features. We took this as evidence to suggest that the vPOp is involved in phonetic encoding and articulatory code generation. However, it is possible that those findings are confounded by the presence of a delay period in the experimental task used. The questions that we would like to ask in the current study are: Is Broca's area really necessary for phonetic encoding? Are the effects that we observed related to verbal working memory or phonetic encoding? To answer these questions we performed another event-related fMRI experiment on the same group of subjects and asked them to perform a modified version of the phonological repetition task, this time with no delay between stimulus presentation and subject response. If the LIFG is involved in phonetic encoding and articulatory code generation, it would show a significant effect of sub-lexical frequency during prompt response trials. Contrary to our expectations, only a region in the left precentral gyrus showed a significant main effect of sub-lexical frequency. We did not find any significant LIFG activity for low vs. high sub-lexical frequency pseudowords. After a close examination of the data, we concluded that the LIFG is not necessary for phonetic encoding and the ventral pars opercularis could not be considered as the site of the mental syllabary. The function of this area seems to be tied to verbal working memory processes. We further discuss our findings with respect to the mechanisms of phonetic encoding and the generation of articulatory codes.

5.1 Experimental Hypothesis

Broca's area has been associated with many different functions related to language processing and speech. In the previous chapter we discussed the role of this area with respect to phonological and phonetic encoding and compared the recent neuroanatomical models proposed by Indefrey and Levelt (2004) and Hickok and Poeppel (2004). Based on the previous study it seems that the ventral part of Broca's area is involved in the process of phonetic encoding and its sensitivity to sub-lexical frequency suggests that it could be the location of the mental syllabary or speech sound map as it has been previously proposed. However, results from the previous study could also reflect a process that is related to working memory rather than phonetic encoding itself. For this study, we were interested in seeing whether we could replicate the findings of the previous study in the absence of a delay period.

The role of Broca's area in verbal working memory (vWM) has been highlighted in the work of Baddeley and the theory behind the phonological loop (Baddeley 1992). According to Baddeley's model of vWM and the phonological loop, this system consists of two subcomponents, a temporary storage component and a sub-vocal rehearsal one. The first subcomponent, the temporary storage is responsible for holding memory traces over a period of a few seconds. During this period the traces decay, unless they are refreshed by the second component, the sub-vocal rehearsal. Based on the findings from studies on neurological patients (for more details refer to chapter 1), it seems that this system uses a type of phonetic information about the items held in the temporary storage component, to help maintain them in memory (Caplan and Waters 1995). Further studies of patients with lesions resulting in phonological loop deficits and neuroimaging studies support the hypothesis of separate storage and rehearsal system with distinct neuroanatomical substrates. While prefrontal BA40 has been associated with storage, Broca's area (BA44) and the premotor cortex (BA6) are thought to be specifically involved in sub-vocal rehearsal

(Vallar and Papagno 2002; Baddeley 2003).

A role for Broca's area in vWM has also been proposed by many other researchers. Neuroimaging studies have shown that this area shows sustained activation during the delay period in delayed serial recall tasks (Chein and Fiez 2001; Strand et al. 2008). A functional segregation of the region between ventral and dorsal has also been proposed, in agreement to the results of our previous study (Chein and Fiez 2001; Chein et al. 2002). The ventral part of the LIFG shows sustained activity during the delay and is sensitive to sub-lexical phonological processes possibly related to sub-vocal rehearsal, while the dorsal part shows a significant decline in activation during the delay period.

A question of interest is whether a role of the ventral part of Broca's area in vWM and sub-vocal rehearsal is consistent with a role in phonetic encoding. At first the two processes (sub-vocal rehearsal and phonetic encoding) may seem distinct. However, according to Hickok and Poeppel (2000; 2004), it is also possible that the two processes are actually using the same cognitive mechanisms. They suggested that vWM relies on an auditory integration network and in this sense it is just another case of sensory-motor integration (Wilson 2001). In Baddeley's model the phonological loop is basically a mechanism to maintain sensory-based representations by means of sub-vocal rehearsal, i.e. using a motor planning system. Based on their proposed model, regions in the STG support the storage of phonological information, while frontal regions, e.g. Broca's area and a dorsal premotor area, support articulatory-based representations. This hypothesis is in agreement with a potential role of Broca's area in imitation and sensory-motor mapping (Iacoboni et al. 1999; Molnar-Szakacs et al. 2005; Nishitani et al. 2005).

However, a recent paper by Makuuchi (2005) challenges the idea that Broca's area is necessary for imitation. Many studies investigating imitation seem to be confounded

by the fact that the actions to be imitated are very simple and repeated many times throughout the experiment. In this sense the presented actions are cues that trigger the execution of already learned actions. Makuuchi then argued that in order to prove that Broca's area is involved in imitation *per se* and not just delayed execution, more complicated, novel actions and in greater variety should be presented as stimuli, so that the subject would be forced to perform a visuomotor transformation on every trial. In an fMRI study manipulating instruction (i.e. the action should be imitated or performed after symbolic instructions) and execution timing (prompt or delayed), it was shown that Broca's area shows a main effect of execution timing, but not instruction. These results were in agreement with studies that have stressed the role of Broca's area in vWM.

In the light of our previous study presented in chapter 4, we also wanted to address this issue and examine the role of Broca's area, and BA44 in particular, in phonetic encoding independent of vWM. If the posterior part of Broca’s area is involved in the process of phonetic encoding, it should show a significant effect for sub-lexical frequency manipulations during a prompt response task. On the other hand, if the area is not involved in syllable articulatory code production, we expect that the effect will not be significant. As discussed in chapters 1 and 4, based on the theory of the mental syllabary, we expect that frequently used syllables (high frequency of occurrence) would be pre-compiled and stored in the area, while infrequent ones (low frequency) would need to be compiled on-line based on their segmental features (i.e. phonemes) independent of whether the task involves a delay or not. We would therefore expect to see higher activation for low vs. high sub-lexical frequency syllables in Broca's area.

To examine these questions we used event-related fMRI to monitor the changes in blood oxygenation while subjects performed a prompt phonological word repetition task. A pseudoword was presented auditorily and it was immediately followed by an auditory probe that indicated whether they should repeat the presented pseudoword

overtly or covertly. The experiment was a modification of the previously performed delayed phonological repetition task (described in chapter 4). It included the same three conditions, length (four vs. two syllables), sub-lexical frequency (low vs. high sub-lexical frequency) and response type (overt vs. covert). The only difference was that for this session, the response occurred immediately after the stimulus presentation. The same group of subjects was also used and the data were acquired on the same day as the delayed phonological repetition task. This also facilitated statistical comparisons between the two tasks (delayed vs. prompt). The stimuli used in this study were different than the ones used in the previous experiment, but with similar features.

We anticipated that by comparing four vs. two syllable pseudowords and low vs. high sub-lexical frequency items during a prompt response task we would be able to identify the areas involved in phonological and phonetic encoding and independent of working memory related processes. As previously mentioned, if the ventral part of Broca's area is involved in phonetic encoding, it would show a main effect of both length and sub-lexical frequency even during prompt response trials.

5.2 Methods

5.2.1 Data Acquisition

Fifteen healthy, right-handed, monolingual native speakers of American English were chosen to participate in the study (8 males, 7 females) with mean age 26 years (range=20-35). This was the same group of subjects that was used for the study presented in chapter 4. However, the results for the two studies were processed separately. Three subjects (two female, one male) were excluded from the analysis because of excessive head motion. During the experimental session, the subjects laid

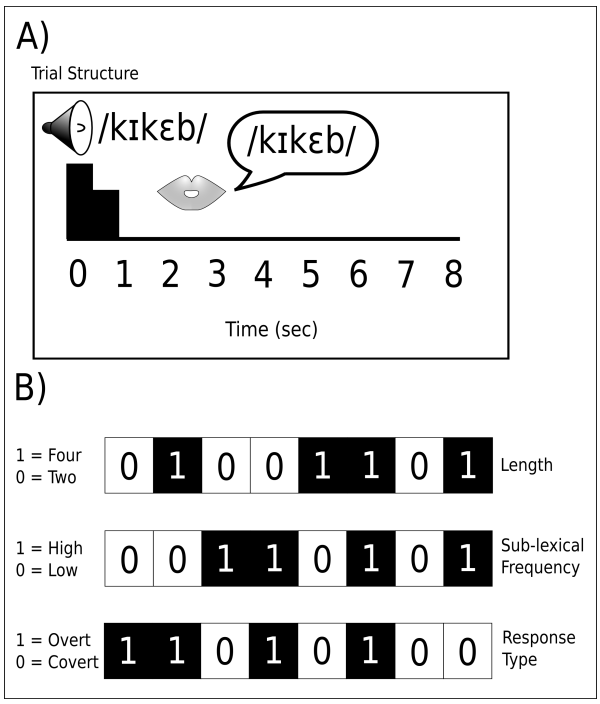

Figure 13: During the experiment, subjects were asked to listen to pseudowords and to repeat them immediately either overtly or covertly. The structure of each trial is shown in (A). The stimulus is presented auditorily at 0s and subjects then wait for the response probe, which comes immediately after the end of the pseudoword. The type of stimulus that will be presented in each trial is determined pseudo-randomly by a combination of 3 m-sequences. In (B) we present an example of 3 binary sequences that resemble those used in the experiment. Each sequence is associated with an experimental factor. In the example provided, the top sequence controls the length of the stimulus (1 for four syllables, 0 for two syllables), the middle sequence controls sub-lexical frequency (1 for high, 0 for low) and the bottom sequence controls response type (1 for overt, 0 for covert).

in the fMRI scanner and were asked to perform a pseudoword repetition task. The presented pseudowords belonged to one of four experimental conditions: four-syllable low frequency, four-syllable high frequency, two-syllable low frequency and two-syllable high frequency.

Over the course of two experimental fMRI runs, subjects were presented auditorily with thirty-six items per condition (for a total of 144 items over all conditions). Immediately after the presentation of the stimulus, a probe (two versions of a bell sound) was heard instructing the subject to repeat the presented word either overtly or covertly (depending on the type of probe; see Figure 13-A). The subjects were given specific instructions to respond as fast as they could upon hearing the probe and were not aware prior to the presentation of the relevant probe whether they would be asked to respond overtly or covertly. Each trial lasted 8 seconds. The length of the trials was chosen to be equal to the trial length of the previous fMRI session

involving a delayed version of the same experimental protocol, to facilitate with between session comparisons.

Stimulus presentation was in a pseudo-random, fast event-related fashion whereby the order of occurrence for the conditions was controlled by a combination of three shifted versions of a binary, 63-bin m-sequence (one shifted by 9 bins and the other by 18 bins with respect to the first one; Figure 13-B). As described in chapter 2, the use of m-sequences ensured that the experimental conditions would be orthogonal to one another and counterbalanced (Kellman et al. 2003; Buracas and Boynton 2002). The binary m-sequence was padded in the beginning with 9 more trials (for a total of 72 trials), which were not analysed for the purposes of this study (please refer to chapter 2 for more on the experimental design).

Since the data for this study were acquired during the same experimental session as the data for the study presented in chapter 4 and because we were also planning to perform comparisons between the two studies, the order of presentation of the studies was counterbalanced. In 8 out of 15 subjects, the prompt response task was performed first, followed by the delayed response task. The subject preparation and image quality control protocols followed are described in more detail in section 4.2.1 There was also no difference between the two studies with respect to the image acquisition protocol and the acquisition of behavioural responses. Please refer to section 4.2.1 for a detailed description of the relevant parameters.

5.2.2 Data Preprocessing and Analysis

All analyses and image preprocessing were carried out using the SPM5 software package and associated toolboxes (http://www.fil.ion.ucl.ac.uk/spm/software/spm5). The preprocessing protocol followed is described in more detail in chapter 3, section 3.2.1. Briefly, images were slice-timing and head motion corrected, registered and

transformed to the MNI anatomical image and finally smoothed with an an isotropic Gaussian filter kernel of 6mm.

Examination of the subject movement parameters provided by SPM5 after motion correction (three factor ANOVA with within-subject factors response type, stimulus length and sub-lexical frequency) revealed a significant main effect of response type in all directions (y, z, roll, pitch, yaw), except translation x (for the five directions $F(1, 11) > 23$, $p < 0.001$). As in the previous study reported in chapter 4 and in agreement with other studies (Barch et al. 1999; Shuster and Lemieux 2005), the incremental movement was overall quite small and greater for overt response trials (mean ±std displacement was 0.039mm ±0.016 for y and z translations and 0.038º ±0.012 for all rotations) than covert response ones (mean ±std was 0.02mm ±0.011 for y and z translations and 0.021º ±0.007 for all rotations). For pitch translation the type of response interacted with the length of the pseudoword ($F(1, 11) = 5.5$, $p < 0.04$). Four-syllable pseudowords also caused greater movement in this direction during overt responses (e.g. mean ±std was 0.054º ±0.016 for four-syllable items during overt response vs. 0.051º ±0.016 for two-syllable items during overt response).

Additional significant effects were present for sub-lexical frequency in the yaw rotation ($F(1, 11) = 19.8$, $p < 0.002$). Low frequency pseudowords (mean yaw displacement was 0.024º ±0.012) produced greater movement than high frequency pseudowords (mean was 0.021º ±0.010).

As described in chapter 3, to remove effects related to subject movement we included the realignment parameters in the design matrix as effects of no interest. In addition, we also added a regressor for images that showed changes in the global signal greater than 1.5% of the mean followed by a greater than 0.5mm incremental movement (Mazaika et al. 2007).

Finally, we inspected the movement parameters for extreme incremental or

absolution motion and excluded from the analysis subjects that did not meet our inclusion criteria as described in section 3.2.2, i.e. absolute motion greater than the voxel size and incremental motion greater than 1mm in translations and 1º in rotations. All subjects met the absolute motion inclusion criteria, but not the incremental motion. Three subjects showed movement greater than our criteria and were consequently excluded from the analysis.

Statistical analysis of the factorial event-related experiment was performed in SPM5 using the FIR approach using a window of 24s. Hypotheses were evaluated using a 3-way, within-subject ANOVA with factors length (four- vs. two-syllable pseudowords), sub-lexical frequency (low vs. high) and response type (overt vs. covert). T-contrasts testing for the predicted shape of the HRF were performed to produce maximum intensity projections (MIP) for the evaluated contrasts. Only one HRF was used which modelled the stimulus presentation and subject response and peaked between 3-7sec. The response type condition was used as a localizer to allow us to define an independent region of interest (ROI) within the left inferior frontal gyrus (LIFG). Statistical parametric maps (SPM) were thresholded at $p < 0.001$ uncorrected at the voxel level and $p < 0.05$ corrected for family-wise error (FWE) at the cluster level (clusters had on average more than 85 voxels; Hayasaka and Nichols 2003).

We performed an additional ROI analysis to examine the effects of working memory on the activation of Broca's area and compared across studies. We defined an independent ROI mask cytoarchitectonically using a map of left hemisphere BA44 (Eickhoff et al. 2005). We were specifically interested in the sensitivity of the LIFG in phonetic encoding and whether this effect is dependent on vWM demands. Because we used the same group of subjects in both experiments (delayed and prompt response), we performed a paired t-test comparing the size of the sub-lexical frequency effect in the LIFG across the tasks. The ROI analysis was performed using the Marsbar SPM5 toolbox (Brett et al. 2002).

5.3 Results

5.3.1 Behavioural Results

To test for effects of length or frequency on subject performance we measured subject response accuracy. Based on previous results, we expected to find a decrease in response accuracy for low frequency pseudowords, but we did not expect to find an effect of length. We performed a 2-way ANOVA with within-subject factors: length and sub-lexical frequency. In agreement to our expectations, we found that there was a significant main effect of sub-lexical frequency only ($F(1,11) = 50.1$, $p < 0.001$). No other effect or interaction was significant. Mean (±std) accuracy rates were 67% (±15) for low frequency pseudowords and 80% (±9) for high frequency pseudowords. All subjects performed with accuracy within three standard deviations of the group mean (74% ±11).

Finally, to ensure that there is a significant difference in sub-lexical frequency between the responses, we calculated the phoneme and biphone phonotactic probability (PP) of the subject overt responses and performed a t-test to compare high vs. low frequency responses. For both biphone and phoneme measurements, the differences were significant ($t(11) = 17.97$, $p < 0.001$ for biphones and $t(11) = 25.30$, $p < 0.001$ for phonemes). Mean (±se) PP for high frequency responses was 0.0206 (±0.0006) for biphones and 0.3838 (±0.0050) for phonemes. Mean (±se) PP for low frequency ones was 0.0025 (±0.0006) for biphones and (0.1165 ±0.0084) for phonemes. From the above results, we can conclude that the subjects perceived the differences between low and high frequency targets and performed the task according to the instructions.

5.3.2 FMRI Results

5.3.2.1 Phonological Encoding

To map the areas involved in phonological encoding we compared the activation levels invoked for processing four- vs. two-syllable pseudowords (over both low and high frequency syllables). A significant main effect of length (four- greater than two-syllable stimuli) was observed in a large perisylvian network extending bilaterally across the superior temporal gyrus (STG), the pre-supplementary motor area (pre-SMA) and the precentral gyrus (PrCG) including small portions of the LIFG (cf. Figure 14-A for whole brain results and Figure 14-B for significantly activated voxels within the LIFG). The largest activations were observed in the left hemisphere for a cluster that covered both the PrCG and STG. In particular for the STG, the cluster covered a large portion of the middle and posterior STG including the upper banks of the superior temporal sulcus (STS) and an area in the junction between the parietal and temporal lobe also referred to as Sylvian parieto-temporal area (Spt) (cf. Table 3 for the coordinates of the significantly activated areas).

To identify whether there was any significant activation within the IFG, we performed a small volume correction within the area of BA44 using the cyto-architectonic probability maps provided by the Anatomy toolbox (Eickhoff et al. 2005). We identified a small cluster (22 voxels) anatomically located on the precentral gyrus that was assigned cytoarchitectonically to BA44 with 30% probability, while it could also be part of BA6 with 20% probability. However, the small volume of the cluster, the low cytoarchitectonic probability and its location on the precentral gyrus raise questions about whether we can consider this as true LIFG activation or whether the activation is simply a result of smoothing. In order to disambiguate this matter we would need to perform an ROI analysis on an independently defined mask of the LIFG. These results are presented below, in

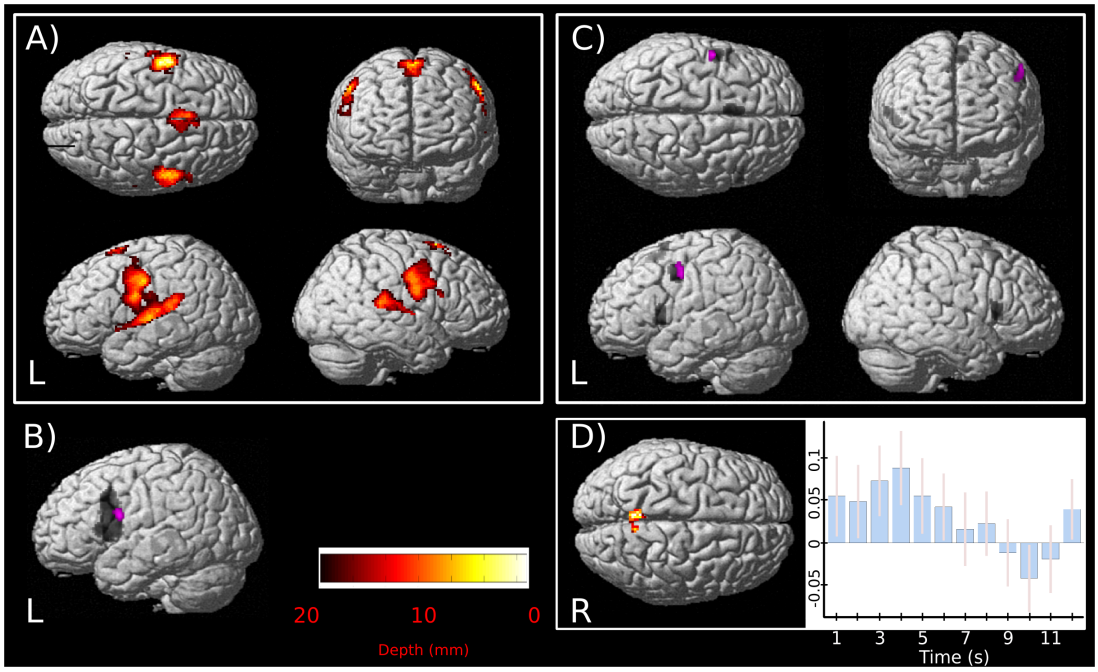

Figure 14: Significant results of the random effects analysis for length (A) and sub-lexical frequency (C). In (A) an extended perisylvian and premotor activation including the LIFG showed significantly higher activation for four vs. two syllables. In (B) we show the main effect of length within left BA44 (significantly activated voxels appear in magenta) after small volume correction (SVC). BA44 (shaded area) was defined using a cytoarchitectonic probability map of the area (Eickhoff et al 2005). Shown in (C) is the main effect of sub-lexical frequency (low vs. high) after SVC. As a mask we used the results previously identified for the same contrast, but during delayed response trials. The mask included the precentral gyrus, the IFG bilaterally and the pre-SMA. Only the cluster in the precentral gyrus (magenta) was significant. Finally, in (D) we present the results for high vs. low sub-lexical frequency. On the left we present a surface rendering of the activation and on the right the contrast estimates for the peak of the activated cluster. Maps are thresholded voxel-wise at $p < 0.001$ uncorrected and cluster-wise at $p < 0.05$ FWE corrected. Colour grading in (A) and (D) reflects depth, with brighter voxels on the surface. The maximum depth of the projected voxels is 20mm. LIFG, left inferior frontal gyrus; L, sagittal view of the left hemisphere; R, sagittal view of the right hemisphere; PrCG, precentral gyrus; pre-SMA, pre-supplementary motor area.

section 5.3.2.3 on the results of LIFG.

Table 3: Brain Regions Modulated by Length and Frequency

Contrast	Region	Coordinates			T	Size
		x	y	z		
Four > Two Syllables	Left precentral gyrus	-54	-4	44	7.12	1839
	* Left superior temporal gyrus	-60	-16	4	6.67	
	* Left sylvian parieto-temporal junction	-52	-34	18	4.98	
	* Left inferior frontal gyrus	-56	4	24	4.13	
	Left pre-supplementary motor area	-2	8	70	5.92	426
	Right superior temporal gyrus	50	-18	8	5.81	551
	* Right sylvian parieto-temporal junction	52	-34	16	5.16	
	Right precentral gyrus	50	-6	26	5.45	958
High > Low Frequency	Left precuneus	-8	-54	54	4.16	169

*Note: Regions significantly activated in the random-effects group analysis ($t(132) > 3.1$, $p < 0.05$ FWE corrected for cluster size). Displayed are the contrasts, the coordinates for the voxels of greatest activity within the activated clusters in MNI stereotaxic space, a description of the region, the t-value and the number of significantly activated voxels. In the case of very large clusters, multiple peak voxels are reported. These are prefixed with a * and they are clustered together with the last non-prefixed entry in the table.*

5.3.2.2 Phonetic Encoding

In order to reveal regions that show an effect of sub-lexical frequency, we compared pseudowords with low vs. high phonotactic probability syllables and segments. Based on our hypothesis, areas that showed a frequency effect reflect the process of phonetic encoding, i.e. articulatory code generation (Indefrey and Levelt 2000). In contrast to our expectations, the whole-brain analysis did not produce any significant results above a threshold of $p < 0.05$ FWE-corrected for cluster-size. Because of the concern that the subject head movement during response might be contributing to increased variability in the data and that we might be suffering from Type II error, we then performed a hypothesis driven analysis. In a previous study involving a delayed version of the current experimental protocol and employing the same subjects, we identified a number of cortical regions that showed a main effect of sub-lexical frequency. In this experiment we are interested in identifying whether these results were task rather than stimulus dependent, i.e. whether they are dependent on the delay period. We therefore created a mask of the significantly activated regions from the previous experimental on phonetic encoding and performed a small volume analysis (SVC). The mask included regions in the left PrCG, the pre-SMA and the IFG bilaterally. The results from the SVC analysis showed a significant activation in the left PrCG (MNI -50 -4 42; $t(132) = 3.86$, $p < 0.02$ FWE-corrected both voxel- and cluster-wise; 23 voxels; see Figure 14-C for significantly activated regions overlaid on the analysis mask). No other region showed a significant effect.

We also looked at the opposite contrast, high vs. low sub-lexical frequency pseudowords, to identify regions that are involved in retrieving pre-compiled articulatory codes from the mental syllabary. There was a significant effect only in the precuneus bilaterally, though left lateralized (see Figure 14-D and Table 3 for the more details). Finally, we looked at the interaction effect for length and sub-lexical frequency, but there were no significant results.

5.3.2.3 Left Inferior Frontal Gyrus

To further test our hypothesis about the involvement of Broca's area in phonetic processing, we performed an ROI analysis. A region corresponding to the LIFG was identified using the contrast overt vs. covert repetition (centre of mass x = -55, y = 8, z = 17, size = 143 voxels; see Figure 15-A). In a random effects two-way ANOVA with factors length (four vs. two syllables) and sub-lexical frequency (low vs. high) the LIFG only showed a main effect of length ($t(11) = 3.6$, $p < 0.003$), but not a main effect of sub-lexical frequency or an interaction effect.

5.3.2.3.1 Delayed vs Prompt Response

Finally, to explore the effect of delay during phonetic encoding on the LIFG, we performed another ROI analysis. The ROI mask was defined cytoarchitectonically for left hemisphere BA44 with centre of mass at x = -53, y = 12, z = 19 and size = 1160 voxels (see Figure 15-B). To compare between the two tasks we performed a paired t-test comparing the size of the sub-lexical frequency effect in the LIFG across tasks. The difference between the two tasks was not significant, though it was only slightly below significance ($t(11) = 1.63$, $p < 0.07$). The contrast values are plotted in Figure 16.

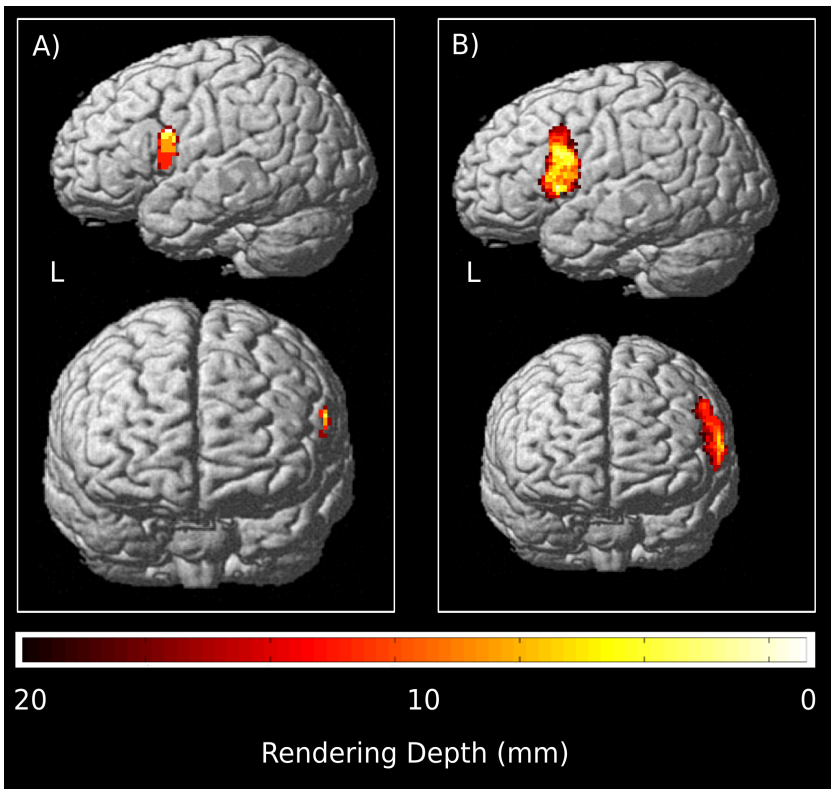

Figure 15: Surface renderings of the masks used for the ROI analyses. Shown in (A) is the result of the SVC analysis for the overt vs. covert responses. As a mask for the analysis we used the cyto-architectonic probability maps for area BA44 (Eickhoff et al. 2005). All significantly activated voxels were included to define a functional mask of the LIFG and used in an ROI analysis examining the effects of length and sub-lexical frequency on the LIFG. In (B) we show the mask that was used in a second ROI analysis contrasting delayed and prompt response trials. This mask was again created based on the cytoarchitectonic probability map for left hemisphere BA44.

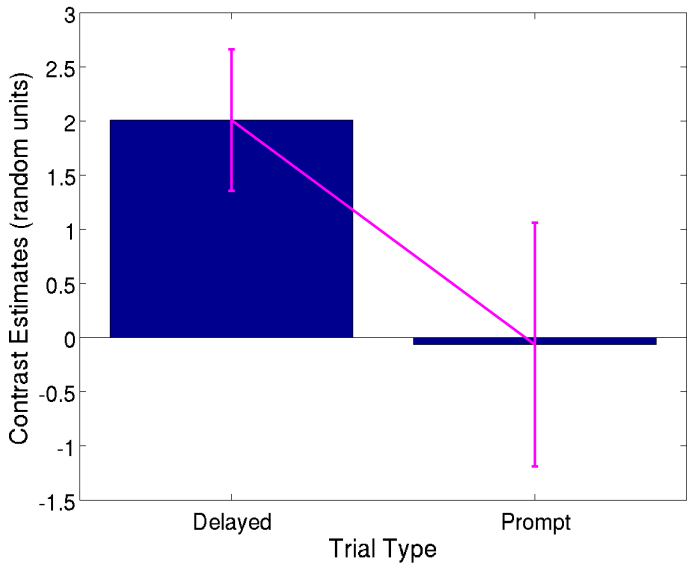

Figure 16: Mean (±se) contrast values for sub-lexical frequency (low vs. high) in the LIFG for the two tasks, delayed and prompt response. The difference in slightly below threshold ($p<0.07$).

5.4 Discussion

In the present study we looked at the role of Broca's area in phonological encoding. A previous study presented in chapter 4 and addressing the same issue, but employing a delayed response task, showed that the LIFG and in particular the ventral pars opercularis is sensitive to sub-lexical features. Based on the predictions from proposed neuro-anatomical models on speech production (Hickok and Poeppel 2000; 2004; Indefrey and Levelt 2000; 2004), these results would imply that the ventral LIFG is involved in phonetic encoding and compiling articulatory codes. However, the results from the delayed response study may have been confounded by the presence of a delay between stimulus presentation and response. Therefore, in the present study we examined the process of phonetic encoding during prompt response trials, avoiding the activation of the phonological loop.

By directly contrasting targets with varying length, we manipulated the load on the system of post-lexical articulatory-motor production and were able to identify a number of key regions underlying articulation and the overall process of transforming phonological word forms to articulatory codes. In summary, these regions included bilateral (although strongly left lateralized) mid and posterior superior temporal and frontal regions (including the dorsal pars opercularis), the premotor cortex and the supplementary motor area. These results are in agreement with current models on word production that describe a left-lateralized, perisylvian network (Indefrey and Levelt 2000; 2004; Hickok and Poeppel 2004; 2007) and the results that we acquired during the previous, delayed phonological repetition study.

By directly contrasting targets with varying sub-lexical frequency, we manipulated the load on the system of phonetic encoding and identified the regions that are involved in the process. Based on the previous neuro-anatomical models of speech production (Indefrey and Levelt 2000; 2004; Hickok and Poeppel 2004; 2007), only regions that are directly involved in phonemic-to-articulatory translation would show an effect for sub-lexical frequency manipulation. Targets with high-frequency

components are processed faster than the ones with less frequent components (Vitevitch and Luce 1998; 2005), which suggests that targets with components of different sub-lexical frequency are processed differently (Guenther et al. 2006). High-frequency clusters are believed to be pre-compiled and their articulatory codes need to be retrieved, while low-frequency clusters need to be compiled on-line on a segment-to-segment basis (Guenther et al. 2006). By directly contrasting low vs. high sub-lexical frequency components, we found significant activation only in a region in the left dorsal pre-motor cortex. This region was in the same location as the premotor cluster identified for the delayed response task. No significant activation was found for the LIFG or any other region.

The absence of significant activation in the LIFG was a surprising result. Even after applying a mask based on the results from the delayed response task, the LIFG cluster was not significant. To further test whether the LIFG had a significant effect of sub-lexical frequency, we performed an ROI analysis. We defined the ROI for the LIFG based on the results for the main effects of response type (overt vs. covert). In the delayed response task, the LIFG ROI analysis showed significant effects for both length and sub-lexical frequency. However, in the prompt response experiment, the LIFG showed a significant main effect of length only. It seems therefore that the LIFG is not necessary for phonetic encoding or articulatory code generation.

So, what is the role of the LIFG? Based on the findings from the delayed response phonological repetition study presented in chapter 4, we concluded that the pars opercularis is functionally segregated into a dorsal and ventral part. The two parts show a different sensitivity to length and sub-lexical frequency, with the ventral part being sensitive to both effects, while the dorsal part was only sensitive to length. In this study, we also observed a main effect of length for the LIFG, in particular for the dorsal part of the pars opercularis, but we did not find any significant activation in the ventral part. We therefore cannot reject the null hypothesis and for this particular experiment an absence of a significant result could mean either that the LIFG is not engaged in the process or that there is no difference in processing low and high sub-

lexical frequency pseudowords (in the LIFG).

To further understand the role of the LIFG in phonetic encoding and vWM, we examined the effect of phonetic encoding with respect to task delay. This analysis did not reveal any significant effects of delay. However, the fact that the results were just below significance implies that there is a trend that the effect of sub-lexical frequency in the LIFG is more active during the delayed response trials. The variance in the contrast effect size for prompt response trials is much greater than that for delayed response trials as shown by the differences in the size of the standard error of the mean. This suggests that there is greater variability within the subject responses in the prompt response trials. This could be caused by subject movement, which affects prompt response trials more, or because the subjects use different strategies to perform the prompt response task, and thus engage Broca's area differently.

In summary, the current results show that the dorsal prefrontal cortex is the only region that shows an effect for both length and sub-lexical frequency. Based on the hypotheses from previous neuroanatomical models on speech processing (Hickok and Poeppel 2004; Indefrey and Levelt 2004), this region would then be the most likely candidate region to be the site of phonetic encoding. This would then mean that the significant sub-lexical frequency effect observed in vPOp during the delayed response trials was related to vWM. These results also suggest that vWM is different than sensory-motor mapping, contrary to what has been claimed by Hickok and Poeppel (2000; 2004). Instead the results are in agreement with what has been observed by Makuuchi (2005) on the role of the LIFG in imitation. As discussed in chapter 1, in a study contrasting the degree of sensory-motor transformation and response delay, he found that the LIFG only showed an effect of delay. Actions that required a prompt response did not significantly engage the LIFG.

In our study we saw that the dorsal part of the pars opercularis showed an effect of length independent of task delay, which could mean that at least the role of dPOp is not limited to vWM related processes, but is more generally involved in motor

planning and syllabification. In a delayed serial recall task, Chein et al. (2002) reported that they also observed a functional segregation of the LIFG in a dorsal and ventral part. The dorsal part exhibited activation that was inverse to the recall success, but correlated with word length, i.e. performance was better and activation was lower when the items were one-syllable words, as compared to three-syllable words. This activation also showed a significant decline over the delay interval, which they took as evidence to mean that the dorsal LIFG is not directly involved in vWM. Based on these results, they concluded that this region contributes to the organization and automation of a sequence of verbal items that will be rehearsed during the delay period. This hypothesis is in agreement with our results and it also follows the hypothesis of Indefrey and Levelt (2000; 2004) and Gelfand and Bookheimer (2003), who also proposed that the LIFG has a role in syllabification or sequencing processes. Gelfand and Bookheimer based their hypothesis on the fact that in a series of sequencing tasks involving hummed notes and strings of syllables the POp did not show an effect of stimulus type, but only an effect of task.

With respect to the ventral part of the pars opercularis, Chein et al. concluded that this region was sensitive to sub-lexical phonological processes related to vWM. In their studies of delayed serial recall, this region exhibited increased activation particularly during the processing of non-words, but also showed sustained activation throughout the delay interval. These findings are in agreement with our study, where we observed an absence of significant activation differences in the LIFG for low vs. high sub-lexical frequency pseudowords. Based on our results we could also add to the Chein et al. hypothesis that the sensitivity of the LIFG in sub-lexical phonological processes during vWM is related to the articulatory codes generated for or during sub-vocal rehearsal in delayed response tasks.

In a review of imaging studies on imitation by Molnar-Szakacs et al. (2005), where they also noted the functional segregation of the pars opercularis, it was suggested that the vPOp is not a premotor region. In particular, they proposed that it is not

directly involved in motor planning, but that it holds some form of representation of the motor plans that is communicated to the posterior part of the superior temporal sulcus. The exact hypothesis claims that the vPOp produces an efferent copy of the target motor plan that is sent to the STS during imitation allowing the prediction of the sensory consequences of planned imitative actions. In this sense the vPOp is not the location of the speech sound map as has been proposed, but it holds a copy of the articulatory codes. The codes themselves are generated elsewhere, which from our data appears to be the dorsal premotor cortex.

Our present findings would be in agreement with such a role of the vPOp. Verbal WM as described by Baddeley (2003) employs a form of articulatory representations. It is therefore possible that a copy of an articulatory representation would be made and stored in the vPOp during tasks with vWM demands or imitation. However, as observed by Makuuchi (2005), the results from many studies on imitation could have been confounded by the fact that the tasks often reminded cued recalls of over-learned actions. The actions that were presented for imitation, such as grasping or making a fist, were often too simple and over-learned for adult humans. As a result, when the subjects were asked (cued) to repeat those actions, they were not necessarily imitating them, but holding them in their working memory and waiting for the cue to execute them. In his study, Makuuchi further showed that there was no difference in the activation of the LIFG when performing a task involving imitation of an action versus a task that simply required the subjects to perform an action following symbolic instructions. The only significant difference observed in the LIFG was when he contrasted delayed vs. prompt response versions of the two tasks. The LIFG was significantly more activated during the delayed response tasks. In this sense, the LIFG is again tied to working memory and is not directly involved in imitation and sensory-motor mapping. The findings from our study are in agreement with these results and further specify the ventral pars opercularis as the portion of the LIFG involved in vWM related processes.

Based on the above results, it is also possible to conclude that the LIFG is not the site

of the mental syllabary, since it shows a dependence on task delay. This role seems to be more suitable for the premotor cortex, in accordance to what had been originally proposed by Indefrey and Levelt (2000). With respect to the LIFG and based on our results on the difference between the dorsal and ventral part of the region, it now seems clearer that the dorsal part is more involved in motor planning processes that could include syllabification and phonological encoding, while the ventral part is involved in vWM processes, possible sub-vocal rehearsal.

Finally, we also wanted to address the results for the opposite contrast, i.e. high vs. low frequency pseudowords, where we only found a significant effect in the precuneus. This result is surprising, considering that the precuneus is not thought to have a dominant role in language production, but has been mostly associated with episodic memory and spatial perception (Cavanna and Trimble 2006). Looking at the contrast estimates (Figure 14-D) this activation does not appear to be a result of noise and we could not think of any reason why there should be a difference in either the spatial perception or episodic memory associations between high vs. low sub-lexical frequency pseudowords. Stimuli were presented binaurally in all conditions and we did not expect any differences in the auditory perception of the stimuli. It could be possible that subjects were using a strategy that could possibly include spatial processing, i.e. visualization, and high frequency pseudowords would be easier to visualise than low frequency ones. However, even if that was the case, it is not clear to us why this would trigger differences related to spatial perception. We currently do not have a sufficient explanation for this result and explaining it could be an interesting direction for future work.

To conclude, in this fMRI study we investigated the processes of phonological-to-articulatory translation and the role of the left inferior frontal gyrus. Based on our findings, we conclude that the dorsal part of the pars opercularis is involved in phonological processing and syllabification, consistent with what we reported in chapter 4. In contrast, the left ventral POp does not seem to be directly involved in phonetic encoding, as previously suggested, and its function should be related to

verbal working memory processes. The dorsal premotor cortex seems to be a better candidate as a site of articulatory code storage for syllables and phonemes. These findings are in agreement with recent observations from the study of imitation and working memory and add more evidence in clarifying the role of the LIFG in vWM and speech production.

Chapter 6: Functional Segregation within Broca's Area: a High Spatial Resolution Approach

In the previous chapters we presented our findings on the role of Broca's area in phonetic encoding and verbal working memory. We showed that there is a functional segregation within the LIFG, following a ventral-dorsal gradient. We also showed that under conditions that would activate the phonological loop, the ventral part of the LIFG, dubbed vPOp, is sensitive to aspects of phonetic processing. On the other hand, the dorsal part of the LIFG, dubbed dPOp, is consistently more sensitive to effects of target length and general phonological processing during both delayed and prompt response tasks. In this chapter, we replicate the results of those studies, using higher spatial resolution images and focusing specifically on the LIFG. Our aim was to validate our previous assumptions about the role of the LIFG and also about the region's functional segregation. We performed the same experiments as described in the previous chapters, but on a different group of subjects. The results from the analysis validate the findings presented in chapters 4 and 5. We also provide further validation of the functional segregation within the LIFG and in particular left hemisphere BA44. We conclude that the results presented in the previous chapters hold and that the LIFG is functionally segregated in a dorsal-ventral orientation. While the dorsal part might be more involved in aspects of motor planning and sequence processing, the ventral part seems to be involved in verbal working memory related processes and possibly in maintaining an active representation of the target stimulus.

6.1 Experimental Hypothesis

With the advent of functional magnetic resonance imaging (fMRI) it became possible for researchers to perform more detailed studies on the functional anatomy of the cerebral cortex. Compared to other non-invasive imaging modalities such as positron emission tomography (PET), fMRI offered, among other things, improved spatial resolution. It became possible to study not just the behaviour of a region as whole, but also the behaviour of different parts within the region, as in the case of regions that are functionally segregated (for examples see Chein et al. 2002; Devlin et al. 2003). More recently it also became possible to study the multi-voxel activation patterns that arise as a result of behavioural stimulation (Cheng et al. 2001; Beauchamp et al. 2004; Kriegeskorte et al. 2006).

As discussed in chapter 2, where we talked about the physical and physiological principles that underlie fMRI, the BOLD signal arises from changes in blood oxygenation. Thus, its origin is in the capillary bed, near the arteries where the oxygenation change occurs, but also in downstream draining veins (Nencka and Rowe 2007). The capillaries are spatially close to the site of neuronal activation, but for magnetic fields less than 3T most of the signal originates in the arterioles as well as much larger structures like the arteries and the draining veins. The larger the structure that generates the signal, the further away the signal is from the source. This is particularly true for large draining veins, where the signal can be displaced by at least a few millimetres with respect to the source of the activation (Lai et al. 1999). Despite all this, the fMRI signal has been shown to be fairly accurate in mapping the areas of activation, at least when it comes to larger anatomical structures and regions. Under higher magnetic fields (> 7 T), where the signal from the capillaries is stronger, it has also been possible to map anatomical structures as fine as the ocular dominance columns (Menon et al. 1997; Cheng et al. 2001).

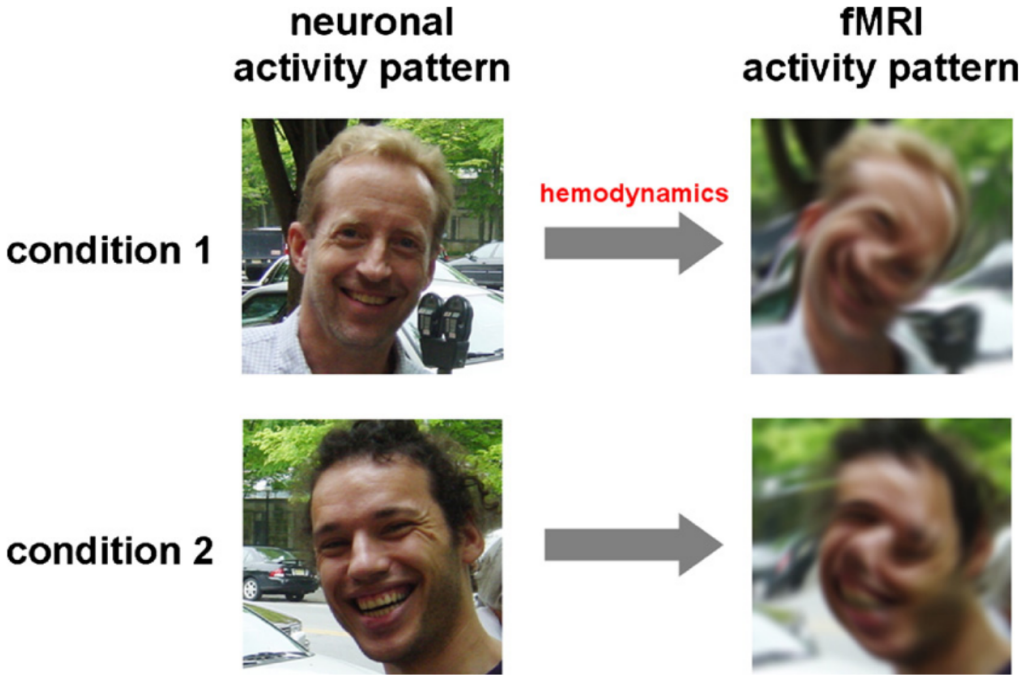

Figure 17: In this figure photographs represent activity patterns. The photographs on the left represent the underlying neuronal activity. The photographs on the right show the same signal filtered through a hypothetical HRF and recorded as an fMRI signal. Despite blurring, distortion and displacement, fMRI activity patterns may distinguish experimental conditions. Even if the neuronal activity pattern is corrupted beyond recognition in the fMRI pattern, information distinguishing the experimental conditions will still be present, as long as the fMRI patterns are replicable and distinct for each condition. Figure adapted from Kriegeskorte and Bandettini 2007.

However, this implies that the focus of the activation can appear displaced. This assumption becomes particularly important when one is interested in studying small anatomical structures or make particular claims about the focus of a particular fMRI activation site. As has been pointed out in many studies, the BOLD image that we are receiving is but a blurred, displaced and distorted image of the underlying neuronal activity (see Figure 17 for a cartoon of the situation; (Kriegeskorte and Bandettini 2007). That is not to say that it does not include valuable information about the neuronal activity itself or that it is not possible to distinguish between conditions

based on the fMRI activity patterns. It is only to raise caution about the assumptions made about the actual site of the activation and to also increase the awareness about the importance and significance of replication. Replication becomes particularly important when the experimental assumptions are tied to fine anatomical differences in the scale of a few millimetres. In such cases, the test-retest reliability needs to be examined to ensure that the anatomical differences observed can be replicated and are therefore caused by the experimental manipulation rather than signal blurring (for examples of studies on fine-grained anatomical differences see Cheng et al. (2001) and Beauchamp et al. (2004)).

In the case of the experiments that we described in chapters 4 and 5, we used a more sensitive than usual technique and thinner than usual slices to address experimental questions on the function and anatomy of the LIFG. We were particularly interested in seeing whether we could observe a functional segregation within the LIFG and in particular the pars opercularis, as has been reported in previous fMRI studies (Chein et al. 2002; Molnar-Szakacs et al. 2005). Even though the question of functional heterogeneity within the LIFG has been addressed quite extensively in the macro-structure, i.e. for the differences between the three anatomical parts of the LIFG, there have been very few studies addressing the question of functional segregation within the parts themselves and in particular the pars opercularis. To our knowledge only three studies have been published so far on this issue (Chein and Fiez 2001; Chein et al. 2002; Molnar-Szakacs et al. 2005). It should also be noted that two of these studies, Chein and Fiez 2001 and Molnar-Szakacs et al. 2005, were meta-analyses of previous studies on verbal working memory and imitation. Thus, the functional segregation did not arise as a result of a direct functional contrast.

In the results presented in chapter 4 we were able to provide evidence for a functional segregation within left hemisphere BA44 in a dorsal and ventral region during a delayed phonological repetition task. These results were extended in chapter 5 by disambiguating the contribution of the delay period. However, because of the

concern that the functional segregation observed could be a result of smoothing, low spatial resolution or displacement, we further wanted to test the validity of these results by following a test-retest reliability check. Therefore, in this chapter we replicated the studies presented in chapters 4 and 5 using a different group of subjects and with a slightly different imaging protocol, whereby the voxel size was reduced by a factor of 6. The key point was to replicate the dorsal-ventral functional segregation within BA44. If we are able to see the same pattern of activation after analysing the new data, then we would be more certain that this segregation is real and not an artifact. In agreement with our expectations, the results that we obtained from the analysis of the two studies presented in this chapter provide further support for the findings provided in chapters 4 and 5 concerning the existence of a functional segregation within left hemisphere BA44.

6.2 Methods

6.2.1 Data Acquisition

6.2.1.1 Delayed Response Experiment

Ten healthy, right-handed, monolingual native speakers of American English were chosen to participate in the study (5 males, 5 females) with mean age 23 years (range = 20-25). The subjects lay in the fMRI scanner and were asked to perform a delayed, pseudoword repetition task as described in section 4.2. This study is an exact replication of the fMRI study presented in chapter 4 with the exception of the scanning protocol. The same stimuli and experimental protocol was used, but the scanning protocol was modified to acquire functional images with higher spatial resolution. The exact details of the study will be described in the following sections.

6.2.1.2 Prompt Response Experiment

The same group of subjects used in the delayed response study described above also participated in the prompt response study and the images for both tasks were acquired during the same scanning session. This study is also a replication of a previous study, presented in chapter 5. Once again, the subjects lay in the fMRI scanner and were asked to perform a prompt pseudoword repetition task as described in section 5.2.

6.2.1.3 Artifacts and Subject Exclusion

Because in this study we acquired images with substantially higher spatial resolution than average (the sides of the voxels were 1.3mm), the images also had lower signal-to-noise ratio (SNR) and were more prone to artifacts during image acquisition. As a result we noted the presence of artifacts in some of the images and had to exclude three subjects (two female, one male) from the analysis. An additional subject (female) was also excluded from the analysis because of problems during the preprocessing of the fMRI data (see Methods section for more details). As a result data from only six out of ten subjects were analysed.

6.2.1.4 Scanning Protocol

Imaging was performed on a 3.0T MRI system using single-shot rate-2 SENSE EPI (de Zwart et al. 2002) as reported in the previous studies presented in chapters 4 and 5. The TR (2sec) and most other scanning parameters apart from the TE, the FOV and the slice thickness were also similar to those reported in chapters 4 and 5. Because of the higher spatial resolution of the images acquired, the SNR was lower in this study compared to the ones reported in the previous chapters. In order to counterbalance this we increased the TE value (TE=35ms) to improve the SNR. The TE value was chosen experimentally after a number of pilot scans. For each image, a total of 20 oblique, axial slices were acquired interleaved (slice thickness = 1.1mm, gap = 0.2mm) with an in-plane resolution of 1.3x1.3mm^2 (144x112 matrix, 18.7x14.6 cm^2 FOV). Four volumes were acquired during each trial. For this study we were only interested in the LIFG. Therefore, we imaged only a limited part of the cortex, focusing on the LIFG. To assist with the selection of the slices, sagittal anatomical images of the lateral view of the left hemisphere were acquired prior to the onset of the experiment. We defined the LIFG by inspection of the major anatomical landmarks that surround the area, such as the Sylvian fissure (ventrally), the vertical ramus of the Sylvian fissure (rostrally), the precentral and the inferior frontal sulci (caudally and dorsally respectively; more details on the anatomy of Broca's area are provided in section 1.2.1). An image of the acquired volume is also provided in Figure 18-C. In Figure 18-B, the partial anatomical image is overlaid on the whole-brain structural image. The red lines show the exact position of the acquired partial volume.

At the end of the scanning session, two high-resolution spin-echo T1 anatomical images were acquired. The first image consisted of the same number of slices and was imaged at the same location as the functional EPI scans (see Figure 18-D). The second image consisted of 30 slices (thicker than the partial volume) and contained the whole brain (see Figure 18-A and B). Both of these images were later used to

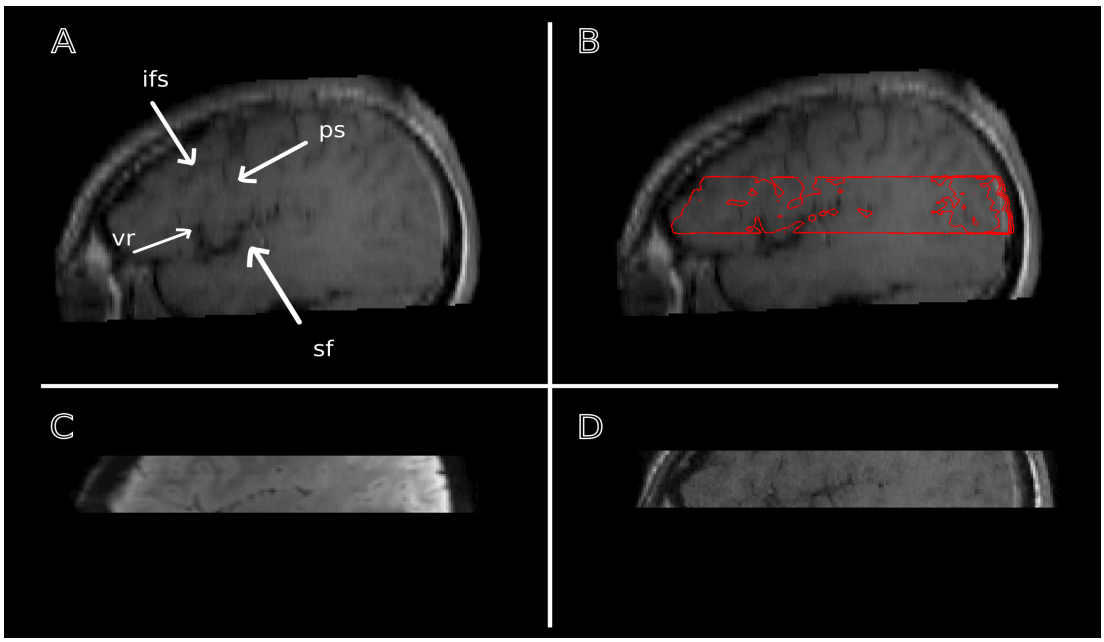

Figure 18: Sagittal views of the acquired volumes (x = -42 in MNI coordinates). Shown in (A) is the left hemisphere of the whole-brain anatomical image acquired at the end of the scanning session. The same sagittal view is also shown in (B), where the partial anatomical (T1) image is overlaid on the whole-brain anatomical image. The outline in red shows the area that was covered by the functional (EPI) and the partial anatomical (T1) images. Major anatomical landmarks (sulci and the Sylvian fissure) are also marked in the overlay. The EPI and partial T1 images are also presented in (C) and (D) respectively. The LIFG was defined anatomically by identifying the Sylvian fissure (sf), the inferior sulcus (ifs), the precentral sulcus (ps) and the vertical ramus of the Sylvian fissure (vr).

facilitate image registration and normalisation. The scanning parameters for the partial volume were: TR=700ms, TE=13ms, 187x187 data matrix with a 22.4x22.4 cm^2 FOV (voxel size 1.2x1.2 mm^2) and 1.1mm slice thickness (with 0.2mm gap). The parameters for the whole brain volume were: TR=700ms, TE=13ms, 256x256 data matrix with a 18.7x18.7 cm^2 FOV (voxel size 0.7x0.7 mm^2) and 3.3mm slice thickness (with 0.2mm gap).

Because for this study we were using very thin slices and we were interested in a very small part of the cortex, it was necessary to ensure that the subjects did not change the position of their head while in the scanner. For this reason, when

positioning the subjects in the scanner, we used additional head padding compared to the previous studies presented in chapters 4 and 5. After the subject was positioned in the scanner and all the padding and straps were applied, we further asked the subjects to demonstrate how much they could move their head. If the setup turned out to be too loose the subjects were repositioned in the scanner. As a last step before starting the scanning, an MRI compatible, fibre-optic microphone was mounted on the head coil and in front of the subject's mouth. As in previous studies, subject responses during the scanning were recorded and the behavioural data were used to validate the subject's performance.

6.2.2 Data Preprocessing and Analysis

All analyses and image preprocessing were carried out using the SPM5 software package and associated toolboxes (http://www.fil.ion.ucl.ac.uk/spm/software/spm5). Details on the preprocessing protocol followed are reported in section 3.2.1. The only difference between the protocol used in this study and the previous ones was during image normalisation. Because only a small part of the cortex was imaged, it was more likely that the automatic registration to the MNI anatomical template would fail. To facilitate the process, two anatomical images were acquired at the end of the scanning session. The partial anatomical image was acquired in the exact location and with the same number of slices and slice thickness as the functional images. The whole-brain image was acquired with additional and thicker slices so that it covered the whole cortex. The partial-volume anatomical images and the functional images were first realigned to one another and then both images were realigned to the whole-brain anatomical image. As a final step the images were realigned to the MNI template.

The next step was to segment the partial anatomical image and use the segmentation parameters for the normalisation of the functional images. Because the volume imaged was very small, for some subjects the alignment process had to be manually

corrected. Errors in the alignment of the cortices will cause errors in the segmentation and eventually in normalisation. For all subjects but one the process worked well. For this subject normalisation failed and the normalized volumes appeared distorted even after several attempts to manually correct the error. We eventually had to exclude this subject from the analysis. As a final step in the preprocessing we applied a 6mm isotropic smoothing kernel. This was the same smoothing as applied to the previous studies presented in chapters 4 and 5. This smoothing kernel is higher than necessary, since it has been shown that when using smoothness-dependent thresholding (as in the case of FWE), using a smoothing kernel twice the voxel size is usually sufficient (Mikl et al. 2008). Thus, in our case a 3mm smoothness kernel would have been sufficient. A higher smoothing kernel suggests that the spatial blurring will be more extensive and the details of the functional patterns will also be blurred. However, we chose to use a higher smoothing kernel so that we could replicate the results from the previous studies.

The quality of the data was further checked using the ArtRepair toolbox for SPM5 (Mazaika et al. 2007) and examining the realignment parameters provided by the SPM5 motion correction procedure. We were particularly interested in scan-to-scan (incremental) motion during the task. As in previous studies, we performed a three factor ANOVA with within-subject factors response type, stimulus length and sub-lexical frequency, and dependent variables the six motion estimates for incremental movement. For both prompt and delayed response studies, the analysis revealed a significant main effect of response type in all directions ($F(1, 4) > 9$, $p < 0.04$ in all directions). In agreement with other studies (Barch et al. 1999; Shuster and Lemieux 2005), the incremental movement for both studies (prompt and delayed) was overall quite small and greater for overt response trials (mean ±std displacement was 0.032mm ±0.004 for translations and 0.022º ±0.004 for rotations) than covert response ones (mean ±std was 0.015mm ±0.003 for translations and 0.011º ±0.002 for rotations).

For delayed response trials there were additional significant main effects for length in

the y translation ($F(1, 4) = 12.2$, $p < 0.03$). The production of four-syllable pseudowords caused greater movement in the y direction than two-syllable ones (mean ±std displacement was 0.0147mm ±0.009 for two-syllable pseudowords and 0.0156mm ±0.012 for four-syllable ones). For the prompt response trials there were significant main effects for length (in both x and y translations $F(1, 4) > 8$, $p < 0.05$) and significant interactions between length and sub-lexical frequency (y translation $F(1, 4) = 8.3$, $p < 0.05$) , between length and response type (y translation and yaw rotation $F(1, 4) > 8$, $p < 0.05$), between sub-lexical frequency and response type (yaw rotation $F(1, 4) = 11.2$, $p < 0.03$) and finally between all the factors (x and y translations $F(1, 4) > 13$, $p < 0.03$). In the three directions that showed additional significant effects (x, y and yaw), four-syllable pseudowords (mean displacement was 0.027mm ±0.004, 0.023mm ±0.019 and 0.014º ±0.005 respectively for the three directions) produced greater movement than two syllable pseudowords (mean was 0.025mm ±0.006, 0.016mm ±0.010 and 0.012º ±0.004 respectively for the three directions). These effects were especially pronounced during overt response trials (in y and yaw directions) and low sub-lexical frequency pseudowords.

To remove effects related to subject movement we included the realignment parameters in the design matrix as effects of no interest. We also used the ArtRepair software (Mazaika et al. 2007) to identify images that showed changes in the global signal greater than 1.5% of the mean followed by a greater than 0.5mm incremental movement. We subsequently inspected the time-series images visually to identify whether there are any images that showed evident motion-related artifacts, e.g. stripes. Images that were identified in this manner were excluded from the analysis by including them in the design matrix as an additional regressor of no interest.

Finally, subjects that showed absolute motion greater than the voxel size (1.3mm) and who also showed incremental motion greater than 1mm or 1º in more than one occasions were excluded from the analysis. All subjects met the motion inclusion criteria and were subsequently used in the analysis.

Statistical analysis of the factorial event-related experiment was performed in SPM5 using the FIR approach. Because the purpose of this study was to replicate and extend the results of the lower-resolution fMRI studies that were presented in chapters 4 and 5, we followed a similar approach as described in the respective chapters for the delayed and prompt response studies. In brief, the HRF for each trial was modelled with 12 bins and we performed a 3-way, within-subject ANOVA with factors length (four- vs. two-syllable pseudowords), sub-lexical frequency (low vs. high) and response type (overt vs. covert).

Before proceeding to the group analysis, we inspected the single-subject results. To ensure the quality of the data, we included a final, functional criterion for subject inclusion in the group analysis. There had to be significantly activated voxels within the LIFG for overt vs. covert response trials in the single-subject results. The LIFG was determined cytoarchitectonically using the BA44 maps provided in the Anatomy SPM5 toolbox (Eickhoff et al. 2005). Significance was determined using a lenient threshold of $p < 0.05$ uncorrected at the voxel level. One subject did not meet the criteria and was excluded from the analysis of both studies (prompt and delayed).

To perform group statistics (random effects) the contrast images for each effect and for the five remaining subjects were submitted to an 1-way ANOVA (with 12 levels). As in the previous chapters, t-contrasts testing for the predicted shape of the HRF were performed that produced maximum intensity projections (MIP) and revealed voxels whose differential activity pattern conforms to the shape of the HRF. For the delayed response trials we used two HRF, one to model stimulus presentation and delay and another one to model the response period (delayed by 6s relative to stimulus onset). For the prompt response trials only the first one was used.

Because for these two high spatial-resolution studies, the focus was on the LIFG, we only performed ROI and small volume correction (SVC) analysis on the group data. To define a mask for the ROI analysis we used the results from the main effect of response type (overt vs. covert response). Once again, we defined the LIFG

cytoarchitectonically and used a lenient threshold at $p < 0.05$ uncorrected (voxel level). This allowed us to identify contiguous voxels within the LIFG that would show a significant main effect of response type. We were therefore able to define an independent region of interest (ROI) within the left inferior frontal gyrus (LIFG). The ROI analysis was performed in Marsbar (Brett et al. 2002).

6.3 Results

6.3.1 Behavioural Results

To test for effects of length or frequency on subject performance we measured subject response accuracy. Based on previous results, we expected to find a decrease in response accuracy for low frequency pseudowords, but we did not expect to find an effect of length. We performed a 2-way ANOVA with within-subject factors: length and sub-lexical frequency. As expected, we found that there was a significant main effect of sub-lexical frequency in the delayed response study ($F(1, 4) = 9.9$, $p < 0.04$). The same effect was just below significance in the prompt response study ($F(1, 4) = 6.8$, $p < 0.06$). No other main effects or interactions were significant. For delayed and prompt response studies, mean (±std) accuracy rates for low frequency pseudowords were 72% (±8.3) and 80% (±4.5), while for high frequency pseudowords the rates 79% (±6.4) and 85% (±5.1) respectively. The mean performance accuracy for all subjects was within three standard deviations of the group mean (79% ±5.6).

Finally, to ensure that there was a significant difference in sub-lexical frequency between the responses, we calculated the phoneme and biphone phonotactic probability (PP) of the subject overt responses and performed a t-test to compare high vs. low frequency responses. For both delayed and prompt response trials the differences in biphone and phoneme measurements were significant. For delayed

response trials $t(4) = 30.1$, $p < 0.001$ for biphones and $t(4) = 19.9$, $p < 0.001$ for phonemes. High frequency responses had greater PP (mean ±se biphone PP was 0.0196 ±0.0004, mean phoneme PP was 0.3708 ±0.0081) than low frequency ones (mean biphone PP was 0.0014 ±0.0002, mean phoneme PP was 0.1080 ±0.0071). For prompt response trials $t(4) = 17.2$, $p < 0.001$ for biphones and $t(4) = 36.5$, $p < 0.001$ for phonemes. High frequency responses had greater PP (mean ±se biphone PP was 0.0212 ±0.0009, mean phoneme PP was 0.3874 ±0.0072) than low frequency ones (mean biphone PP was 0.0016 ±0.0003, mean phoneme PP was 0.1072 ±0.0014). From the above results, we can conclude that the subjects perceived the differences between low and high frequency targets and performed the task according to the instructions.

6.3.2 FMRI Results

6.3.2.1 Delayed Response Study

To test our hypothesis about the involvement of Broca's area in phonetic processing, we performed an ROI analysis. A region corresponding to the LIFG was independently identified for each subject using the contrast overt vs. covert repetition as described in section 6.2.2. In a two-way ANOVA with factors length (four vs. two syllables) and sub-lexical frequency (low vs. high) the LIFG showed a main effect for both factors ($F(1, 4) = 18.8$, $p < 0.03$ and $F(1, 4) = 18.6$, $p < 0.03$ for length and frequency respectively). Post-hoc comparisons revealed that four-syllable pseudowords had a greater effect than two-syllable ones ($t(4) = 4.3$, $p < 0.02$) and low sub-lexical frequency pseudowords had a greater effect than high frequency pseudowords ($t(4) = 4.3$, $p < 0.02$). The interaction was not significant ($F(1, 4) = 0.6$, $p < 0.48$). These results replicate those reported in chapter 4.

We then examined the data to see whether there were any signs of functional

segregation within the LIFG and in particular the pars opercularis, as had been observed in other published studies (Molnar-Szakacs et al. 2005) as well as our own, reported in chapter 4. Based on the previous results, the LIFG appeared to be functionally segregated in a dorsal-ventral direction. To examine this effect we performed an SVC of the group data and identified clusters within the LIFG that showed a main effect of length and sub-lexical frequency. Once more, as a mask for the LIFG we used a cytoarchitectonic mask of left hemisphere BA44 (Eickhoff et al. 2005). We then compared the two clusters with respect to the Euclidean distance of the cluster peak and their overlap. The two clusters (shown in Figure 19) were only partly overlapping (4 voxels out of 15 and 200 respectively for the length and sub-lexical frequency clusters). The distance between the cluster peaks was 20 mm, which is greater than the smoothing kernel (6mm) by at least a factor of 3. The cluster showing a greater effect of length occupies the dorsal part of the pars opercularis (cluster peak located at MNI [-60 4 14]), compared to the cluster showing a greater effect of frequency, which extends from the banks of the Sylvian fissure and follows up the vertical ramus of the fissure (cluster peak located at MNI [-46 18 10]). We will refer to the cluster identified during the length condition as dPOp (dorsal pars opercularis) and the cluster identified for the frequency condition as vPOp (ventral pars opercularis), because of their anatomical differences and in agreement with the clusters that we identified in chapter 4.

6.3.2.2 Prompt Response Study

A similar analysis as described above was also applied to the prompt response study. In summary, we extracted a functional mask of the LIFG from the results of a SVC analysis on the main effects of response type (overt vs. covert repetition). In a random effects two-way ANOVA with factors length (four vs. two syllables) and sub-lexical frequency (low vs. high) the LIFG showed a significant main effect only for length, but not for sub-lexical frequency or the interaction between the two factors ($F(1, 4) = 11.3$, $p < 0.03$ for length, $F(1, 4) = 4.8$, $p < 0.10$ for frequency and $F(1, 4)$

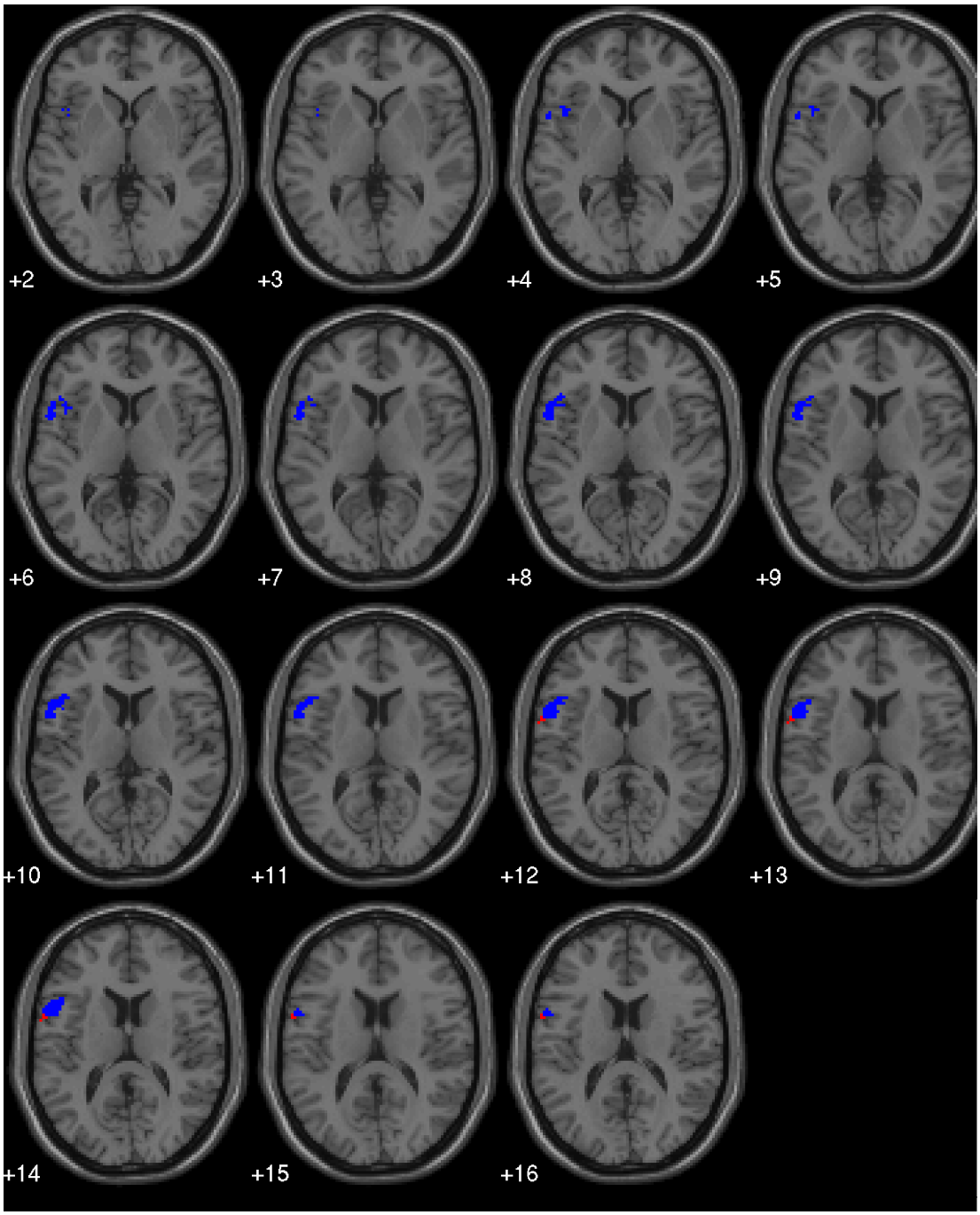

Figure 19: Significant activations within left hemisphere BA44 as defined by a cytoarchitectonic probability map of the area (Eickhoff et al. 2005). Shown in red are significantly activated voxels for four vs. two syllables. This cluster extends from z=12 to z=16. The cluster peak is located dorsally at [-60 4 14]. Shown in blue are significantly activated voxels for low vs. high sub-lexical frequency. This cluster extends from z=2 to z = 16. The cluster peak is located at [-46 18 10]. Activations are thresholded at p < 0.05 uncorrected voxel-wise. Coordinates are in MNI space. Left hemisphere is shown on the left.

= 0.0, $p < 1$ for the interaction). Post-hoc comparisons revealed that four-syllable pseudowords had a greater effect than two-syllable ones ($t(4) = 3.4$, $p < 0.02$). Once again, the results replicate the results that we reported in chapter 5 and show that the LIFG is not necessary for phonetic encoding. The sub-lexical frequency effect observed in the LIFG during the delayed response task is dependent on the presence of a delay period and could probably be attributed to sub-vocal rehearsal.

6.3.2.3 Un-smoothed Data

Following the results from the group analysis on the smoothed data, we also wanted to look at the information that is contained within the un-smoothed data and whether the same analysis would confirm our results. Therefore we repeated the same ROI analysis as reported above, only this time we used the un-smoothed data. Because we were only interested in looking at specific effects, instead of an ANOVA we performed two one-way t-tests looking at the contrasts between four- vs. two-syllable pseudowords and low vs. high sub-lexical frequency. Because in the analysis of un-smoothed data the activation patterns are much more specific to the individual subjects and less overlapping across the group, it is generally not advised to perform group analysis. To overcome this issue, we re-defined the ROI mask used for this analysis based on the single-subject SVC results for the main effect of response type (overt vs. covert repetition). For every subject, an individual ROI mask was created and the contrast values were calculated within that mask. These values were then used for the group analysis.

The results replicated the results presented above for both the delayed and prompt response studies. In summary, for both studies (delayed and prompt response) the processing of four syllable pseudowords produced greater activation in the LIFG than the processing of two-syllable pseudowords ($t(4) = 2.5$, $p < 0.04$ and $t(4) = 2.9$, $p < 0.02$ for delayed and prompt studies respectively). The processing of low vs. high

sub-lexical frequency pseudowords also produced significantly greater activation of the LIFG for both studies ($t(4) = 2.5$, $p < 0.04$ and $t(4) = 2.7$, $p < 0.03$ for delayed and prompt response studies).

The significant difference between low and high frequency pseudowords in the LIFG for the prompt response study is in contrast to the results presented above for the same study, but using the smoothed results. To further assess the accuracy of the results we plotted the time course of the contrast estimates, which would allow us to assess whether the results were because of the influence of noise or actual activation. In Figure 20 we present the contrast estimates over time. The presented time course

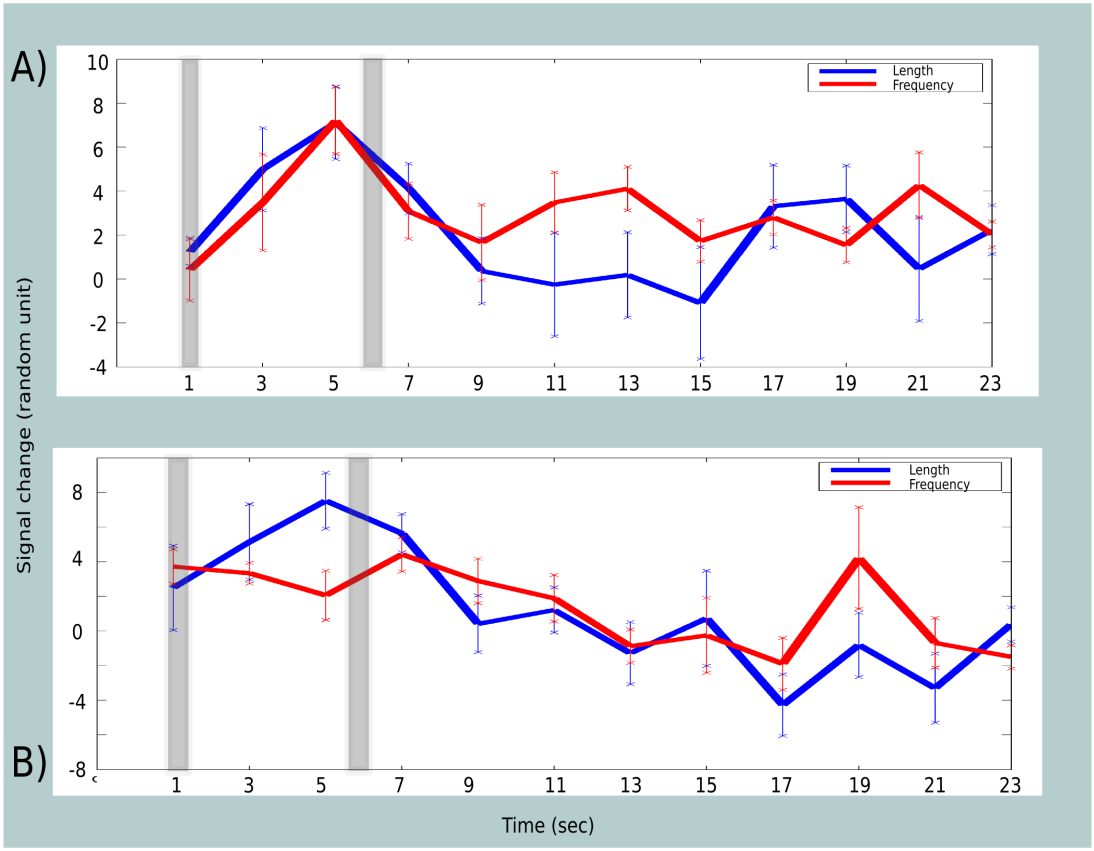

Figure 20: Plots of the FIR contrast estimates over time using un-smoothed data from an ROI analysis of the LIFG. The plot in (A) shows the estimates for the contrast four vs. two syllable pseudowords (length, shown in blue) and low vs. high sub-lexical frequency (frequency, shown in red) for the delayed response trials. The plot in (B) shows the estimates for the same contrasts as in (A) but for the analysis of the prompt response trials (same colour coding). The two grey bars mark the presentation of the stimulus and the response probe respectively.

for the contrast low vs. high sub-lexical frequency for the prompt response study (the red line in Figure 20-B) gives the impression that the result is noise and not actual signal. Based on dynamics of the HRF, we would expect that a plot of the contrast estimates over time would show a peak in the activation around 4-6s and then the activation would drop and return to baseline. Because the contrast estimates presented here are generated by contrasting low vs. high sub-lexical frequency pseudowords and the baseline is rather high, the curve of the contrast estimates is not as smooth as one would expect if contrasting with a low-level baseline such as rest. Still, we can see that the contrast estimates in A and the plot for length in B roughly follow this pattern. They show a peak around 5s and then the signal difference reduces. In the case of the contrast estimates for sub-lexical frequency in B, there is hardly anything resembling a peak around that time, but a rather elevated activation for low vs. high sub-lexical frequency pseudowords, which is rather difficult to explain in terms of the HRF shape. Therefore, we conclude that the effect should be treated with caution, as it is possible that it is not true activation.

6.4 Discussion

In this chapter we replicated the results that we presented in chapters 4 and 5 and verified the role of the LIFG in verbal working memory and sequence processing. By replicating the two studies of delayed and prompt response trials, we were able to show that the dorsal-ventral segregation observed in chapter 4 is real and does not arise as a result of BOLD-related blurring or displacement. In doing so, we were also able to validate the assumptions that we made in the previous studies regarding the role of the LIFG in phonological processing and verbal working memory.

As mentioned in chapters 4 and 5, in recent models of speech production the LIFG has been assigned to different functional roles. According to some models it is involved in phonological processing and syllabification (Indefrey and Levelt 2000;

2004), while others claim that its role is in phonetic encoding and articulatory code generation (Hickok and Poeppel 2000; 2004; 2007). Others again assign the function of the LIFG to the process of verbal working memory (Chein et al. 2002; Baddeley 2003). To address this issue and to disambiguate the role of this region, we conducted a series of fMRI studies examining the neuroanatomical substrates of phonological and phonetic encoding. Based on the theory initially proposed by Indefrey and Levelt (2000), syllables with low frequency components should be compiled on-line when compared to syllables with high frequency components, which should be pre-compiled and retrieved from a cortical area that has been dubbed the mental syllabary. Therefore, by manipulating the sub-lexical frequency of auditorily presented pseudowords we were able to examine the areas that are involved in phonetic encoding and articulatory code generation.

In chapter 4, we showed how the ventral part of BA44 was sensitive to the manipulation of sub-lexical frequency, when the task had working memory demands and engaged the phonological loop. However, in chapter 5 we showed that the same region did not show a significant effect for the same manipulation, but under conditions where we would not expect the phonological loop to be activated. We concluded that the ventral part of the LIFG was not necessary for articulatory code generation and its role could be related to verbal working memory, as it has been claimed by Chein et al. (2002).

In the studies presented in this chapter, we also manipulated the length and sub-lexical frequency of the presented pseudowords under both prompt and delayed response conditions. However, in these studies our focus was constrained to the LIFG and instead of imaging the whole brain, we only focused on a thin belt covering a big part of the LIFG. We then performed an ROI analysis to examine the effects of length and sub-lexical frequency specifically on the LIFG. The results from the ROI analysis replicated the results presented in chapters 4 and 5. It showed that the LIFG is sensitive to both length and sub-lexical frequency during the delayed response trials and in a further SVC analysis, we were also able to show that there

was a functional segregation within left hemisphere BA44. The two clusters that showed significant main effects for length and frequency respectively were located far apart from one another and with very small overlap, thus supporting the assumption that they are distinct functional regions. This brings further evidence in support of the existence of a dorsal-ventral segregation within the LIFG.

The analysis of the second study, which involved prompt response trials, further confirmed the assumptions about the role of the LIFG in verbal working memory. The results from the ROI analysis showed that the LIFG showed a significant main effect of length only and not of sub-lexical frequency, when there were no verbal working memory demands. The results exactly match the results that were presented in chapters 4 and 5.

To further test the validity of these findings and to also check how robust these results would be even under conditions of no data smoothing, we performed a further analysis this time using the un-smoothed data. In an independently identified ROI for the LIFG, we tested for the effects of length and sub-lexical frequency. The analysis of the un-smoothed data again replicated our previous results and confirmed that the LIFG is sensitive to both length and sub-lexical frequency effects under conditions of verbal working memory, while in the prompt response study it is only sensitive to length. With respect to the latter results, even though the results from the ROI analysis of the prompt response study showed that there was actually a significant effect of sub-lexical frequency, a subsequent examination of the contrast estimates suggested that this result was not valid and should be considered as noise. Even though in this study the baseline used for the comparisons is quite high and we would not expect the contrast estimates to show the exact shape of the HRF, we would still expect to see an approximation of the HRF and the peak to be between 4-6s. The contrast estimates presented in Figure 20-B for the effect of sub-lexical frequency do not show any such pattern and they are also in striking contrast to the contrast estimates for the effect of length.

The presence of false activation in the un-smoothed data does not come as a surprise. When compared to smoothed data, un-smoothed data maintain a more accurate representation of the underlying activation patterns. However, the SNR in un-smoothed data is also lower than in smoothed data. In our study this effect became even more pronounced because of the fact that we used a higher than usual spatial resolution (voxel were 1.3mm in each side). Higher spatial resolution means less signal dropout in sensitive cortical areas, but it also means lower SNR. In our case the drop in SNR caused severe problems during image acquisition, which meant that we had to exclude three subjects from the analysis due to the presence of artifacts. What it also means is that the statistical power is lower, which is catastrophically combined with the fact that the increased number of voxels in the data increases the severity of the multiple comparisons problem. To bypass this problem, we only performed ROI and SVC analyses on functionally identified ROI. This approach was anatomically more constrained and well-suited for the purposes of this study. It allowed us to replicate previous results and confirm the validity of the previous claims both about the role of the LIFG in verbal working memory and sequence processing and about the dorsal-ventral functional segregation of BA44.

Even though for the analyses presented in this chapter we have not taken full advantage of the high spatial resolution, we have shown that the high resolution samples that we used in the group study are of good quality. In future work, the high spatial resolution data could potentially be used for an information type of analysis (Kriegeskorte et al. 2006; Kriegeskorte and Bandettini 2007) or single-subject analysis and they could reveal more information about the underlying patterns of activation and their distribution (Haxby et al. 2001; Cheng et al. 2001). Even though caution should be exercised in the interpretation of such results, there is still more information that can be derived. In the activation-based[9] analyses that we have presented for these data, we have provided evidence to support the claim that the ventral and dorsal parts of BA44 code for different types of information when

9) The term activation-based is used in contrast to the term information-based analysis to denote statistical approaches where the focus is on whether one condition activates one region more than another (Kriegeskorte et al. 2006).

processing auditory stimuli. One is more sensitive to length and the other more sensitive to phonetic information during the delay period.

An interesting question that we could ask in the future is about the information that is actually processed in the regions. Activation-based analyses only show the sensitivity of a region to one condition versus another, but they do not reveal anything about the information that the region manipulates. It is still possible that a region is involved in a specific process, even though the effect cannot be identified by means of an activation-based analyses. For example, the response of a region could be positive for one condition and negative for another condition (versus the same baseline). The additive effect of the positive and negative responses would mean that the effect of these conditions on the region would not be significant in an activation-based approach (Kriegeskorte et al. 2006; Kriegeskorte and Bandettini 2007). In our case, the ventral part of the LIFG did not show any significant activation for sub-lexical frequency during prompt response trials, but does this really mean that the region is not at all active during the task, or that there is no difference in the processing of high and low frequency pseudowords during prompt response trials? Or does it simply mean that the effects are cancelling each other out? In order to be able to answer these questions we would need to follow up with an information-based analysis.

In summary, in this chapter we presented the results from the replication of the studies presented in chapters 4 and 5. The results did indeed replicate and provide further support to the claim that the LIFG and in particular, left hemisphere BA44 is functionally segregated into a dorsal and ventral part. Based on our results and the results from previous studies on the role of the LIFG, we conclude that the dorsal part is involved in phonological processing and syllabification, while the ventral part is involved in verbal working memory and in maintaining an active articulatory representation of the target stimulus.

Chapter 7 Discussion

In this thesis we examined the neuroanatomical substrates of phonetic encoding and the generation of articulatory codes with an emphasis on the role of the left inferior frontal gyrus (LIFG). In summary, we have provided evidence to support a role of the premotor cortex in phonetic encoding and articulatory code generation. We have also shown how the LIFG is functionally segregated. Based on the evidence presented in this thesis, the LIFG, and in particular BA44, seems to be segregated following an approximately dorsal-to-ventral gradient. The dorsal part of the posterior LIFG, which we have dubbed dPOp, shows a significant difference in the magnitude and extent of its activation when processing longer vs shorter pseudowords (measured in number of syllables and phonemes). This difference is irrespective of whether the task includes a delay period or not. This evidence suggests that the dPOp is involved in aspects of phonological or motor planning processing. On the other hand, the ventral part of the posterior LIFG, which we have referred to as vPOp, only shows a significant difference in the magnitude and extent of its activation when processing low vs. high sub-lexical frequency pseudowords in the presence of a delay period and during engagement of the phonological loop. This evidence suggests that the posterior LIFG, which traditionally has been referred to as Broca's area, should not be treated as a functionally homogeneous region, in particular when referring to its involvement in the different aspects of phonological processing. We propose that the existing models are revised to include a more complex role for the posterior LIFG and to incorporate evidence of its functional segregation. This concluding chapter will draw together the main findings of the thesis and evaluate them in relation to previous work on phonetic encoding, sensory-motor integration and the role of Broca's area. Possible directions for future work are also outlined.

7.1 Phonetic Encoding and the Generation of Articulatory Codes

The overall aim of this thesis was to study language production and to further characterize the role of the LIFG in language production. A detailed investigation of neuroanatomical models of language production, presented in chapter 1, revealed how there are great differences of opinion between the different models regarding the role of the LIFG. However, these inconsistencies can be addressed experimentally. In the preceding chapters, we addressed some of these issues and here we provide a more detailed view on how our findings could contribute to the revision of some of the models of language production.

In chapter 1, we mentioned how, according to a model of language production proposed by Indefrey and Levelt (2000; 2004), the phonological/phonetic system consists of a phonological encoding process, a phonetic encoding process and the process of articulation (see Figure 1 for a diagram of the proposed phonological/phonetic system). In their model, phonological encoding consists mainly of processes related to segmentation and syllabification, while phonetic encoding is associated with the mechanism of generating articulatory codes. Based on an extended literature review of imaging studies, each of these processes was also associated with a cortical region. Phonological encoding was associated with the posterior LIFG, while phonetic encoding was associated with the left premotor cortex.

To evaluate their theory and test their neuroanatomical hypotheses, we conducted a series of fMRI studies to examine directly the regions that are involved in phonological and phonetic processing. We assumed that regions that are part of the phonological/phonetic system would be sensitive to articulatory load, as this is

expressed in the length of the target item. A subset of these regions would be specifically sensitive to phonetic encoding. To identify these regions we compared low vs. high sub-lexical frequency pseudowords. Based on the proposed theory of phonetic encoding, the two different categories would be processed differently in the brain. Low frequency components are thought to be compiled on-line, while high frequency components are pre-compiled and stored in the mental syllabary. When needed these components are therefore retrieved, rather than compiled. We also assumed that this processing difference between low and high frequency components would be reflected in the magnitude of the region's activation. Compiling the articulatory codes of a target would require more resources than retrieving a set of pre-compiled codes. Our findings partly support the model presented by Indefrey and Levelt.

By comparing four- vs. two-syllable pseudowords, we identified the bilateral posterior, superior temporal gyrus, the bilateral premotor cortex and the left inferior frontal gyrus (BA44) as the underlying regions of the phonological/phonetic system. When the Indefrey and Levelt model (Indefrey and Levelt 2000) was first presented, the superior temporal gyrus was identified as part of the phonological/phonetic system and it was considered as one of the candidate regions to support phonetic encoding. However, this role was later revised (Indefrey and Levelt 2004) and this region was no longer included in the phonological/phonetic system, but assigned to the level of lexical phonological access. Because our experiment consisted of pseudowords that had also been controlled for immediate phonological neighbours, we did not expect that regions related to the processing of lexical information would appear, yet we observed strong bilateral activation along the superior temporal gyrus (STG) with a peak in the left hemisphere posterior STG for four- vs. two-syllable pseudowords. We take this as evidence that the STG is also involved in phonological/ phonetic processing during both perception and production, as has been proposed by others (Hickok and Poeppel 2000; 2004; 2007).

By contrasting low vs. high sub-lexical frequency pseudowords, we were able to identify the regions involved in phonetic encoding. In two experiments, examining the effects of sub-lexical frequency and task delay, we found that the left hemisphere premotor cortex is the only region that is sensitive to the phonetic features of the pseudowords independent of task delay. Therefore, we concluded that the left premotor cortex is involved in phonetic encoding, in agreement with the Indefrey and Levelt model. Where our findings seemingly disagree with the Indefrey and Levelt model is with respect to the role of the LIFG. Based on their model, the LIFG is involved in phonological processing, and in particular syllabification. However, in our experiments we observed that the LIFG is not a homogeneous region and there is a functional difference between the dorsal and ventral part of the left hemisphere BA44. While the dorsal part of the area shows consistent sensitivity to the target length, independent of task delay, the ventral part shows sensitivity to the target's sub-lexical frequency, but only during delayed response trials. In this sense, it is possible that the dorsal part of the LIFG is involved in phonological processing, while the ventral part is involved in verbal working memory. In light of this evidence we propose that the model of lexical production proposed by Indefrey and Levelt should be revised to take into account the functional segregation of the LIFG and also the role of the STG in phonological processing.

Another neuroanatomical model of language processing that also referred to Broca's area as one functional entity is the dual-stream processing model proposed by Hickok and Poeppel (Hickok and Poeppel 2000; 2004; 2007). In this model, phonological and phonetic processing is associated with the dorsal stream of language processing and the process of sensory-motor mapping. The main hypothesis is that acoustic/phonetic speech codes are associated with articulatory-based speech codes through a process of sensory-motor mapping. This process is particularly active during the early years of development and language acquisition, as well as when one is found in a new or unfamiliar linguistic environment. However, it seems to be less active in adulthood and everyday life. This process is similar to the process

of phonetic encoding and the theory of the mental syllabary. The articulatory codes for high sub-lexical frequency targets (i.e. well-rehearsed and frequently used targets) are pre-compiled and stored in the mental syllabary, also referred to as the speech-sound map. On the other hand, low sub-lexical frequency targets (i.e. much less rehearsed and encountered targets) activate the sensory-motor mapping mechanism for the generation of the articulatory codes. In the proposed model, Hickok and Poeppel name the LIFG and the premotor cortex as the storage site for the articulatory codes (the speech-sound map), while the posterior STG is thought to be the interface for sensory-motor mapping.

In partial agreement with the Hickok and Poeppel model, we found that a region in the premotor cortex is sensitive to the phonetic features of the presented pseudowords, which could suggest that this region is involved in sensory-motor mapping. However, we could not identify such a role for the LIFG. As previously mentioned, we observed a functional segregation of the area into a dorsal and ventral part, but neither of the sub-regions could fulfil the criteria of a speech-sound map. For one thing, the dorsal area did not show any significant difference in the processing of low vs. high sub-lexical frequency pseudowords, particularly in the absence of task delay. Furthermore, the ventral part of the area only showed sensitivity to sub-lexical frequency in the presence of task delay. This allows us to conclude that the LIFG is not involved in sensory-motor mapping. If that were the case, we would expect to see some differences in the processing of low vs. high sub-lexical frequency pseudowords independent of task delay, as can be observed in the premotor cortex. Even though null results should generally be treated with caution, the replication of the findings presented in chapters 5 and 6 allows us to have a greater degree of certainty about the validity of these results and the conclusion that the LIFG is involved in verbal working memory processes.

7.2 Verbal Working Memory and the LIFG

A further point of contrast with the Hickok and Poeppel model is their theory on verbal working memory and its anatomical substrates. Based on their theory, verbal working memory and in particular the mechanism of the phonological loop could be considered a special case of sensory-motor mapping. In their own words "... This sensory-motor loop in the dorsal stream provides the functional anatomical basis for verbal working memory, that is, the ability to use articulatory-based processes (rehearsal) to keep auditory-based representations (storage) active" (Hickok and Poeppel 2004). However, our findings point to a different conclusion, i.e. that verbal working memory is indeed different than sensory-motor mapping, at least with respect to their neuroanatomical substrates. We observed that the ventral part of the LIFG showed a significant main effect of sub-lexical frequency only during delayed response trials, but not during prompt response trials. This suggests that there is a difference between verbal working memory and sensory-motor mapping and that the ventral part of the LIFG is functionally involved in verbal working memory, as has been suggested by Chein et al. (Chein and Fiez 2001; Chein et al. 2002), while the premotor cortex is involved in phonetic encoding and sensory-motor mapping.

The model proposed by Hickok and Poeppel is not meant to be a model of verbal working memory as such, but of language processing in general, whereby the authors make a very worthy attempt to find common underlying processes for a lot of the language related functions. To a great extent, they based this common platform on the theory of the mirror neuron system and its extension for language (Rizzolatti and Arbib 1998) and propose that most language processes, including verbal working memory, can be accommodated by a series of transformations, e.g. between acoustic and lexical information or in the case of language production, between acoustic and articulatory codes. As part of this endeavour they also develop a theory on the relationship between sensory-motor mapping and the phonological loop mechanism that has been proposed by Baddeley (1992; 2003).

As Hickok and Poeppel themselves acknowledged, Baddeley himself did not have any type of sensory-motor mapping in his mind, when he was describing the mechanism of the phonological loop, though the two theories are not incompatible. As we mentioned in chapter 1, the phonological loop is thought to be largely dependent on the acoustic/phonological characteristics of the target stimuli (Baddeley 1966), but more specifically it has been linked to the generation of speech-motor plans, i.e. articulatory codes (Caplan and Waters 1995; Baddeley 2003). Regarding the neuroanatomical substrates of the phonological loop, Baddeley suggested that the phonological short term storage is located on the inferior parietal lobe (BA40), while Broca's area (BA44) and the premotor cortex (BA6) support sub-vocal rehearsal. Our findings provide support for the involvement of Broca's area and the premotor cortex in verbal working memory tasks. However, we have also identified that the role of the premotor cortex and the dorsal LIFG is more generally related to language processing and is not specific to verbal working memory. The only region that, based on our findings, seems to be specifically involved in verbal working memory and possibly sub-vocal rehearsal processes is the ventral LIFG.

7.3 The Functional Segregation of BA44

As already mentioned in chapter 4, a role of the ventral LIFG (vPOp) in verbal working memory is not inconsistent with other neuroimaging studies. In two imaging studies on verbal working memory, Chein et al. were the first to observe distinct patterns of activity within two subregions of the LIFG (Chein and Fiez 2001; Chein et al. 2002). The ventral part was sensitive to lexical status (greater activation for non-words vs. words) and sub-lexical phonological processes, while the more dorsal one tracked with recall performance and it was thought to be involved in sequence processing. Even though in those studies, the location of the LIFG foci was not specified using cytoarchitectonic maps, a later examination of the cluster peaks

identified that both the dorsal and the ventral foci reported in the above studies are within left hemisphere BA44. This is the exact pattern that we observed in our own study, i.e. that BA44 is functionally segregated in a dorsal and ventral part.

To our knowledge there are no anatomical data to support the existence of a functional segregation within BA44. As we have mentioned in chapter 1, BA44 is a dysgranular region, i.e. layer IV is not very clearly delineated, and cytoarchitectonically it seems to be a transition area between premotor BA6 and prefrontal BA45. The dorsal part of BA44 is neighbouring BA6, while the ventral part of BA44 borders BA45. This relation can also be observed in the cytoarchitectonic probability maps, where parts of the dorsal part of BA44 overlap with BA6, while parts of ventral BA44 overlap with BA45. A similar relationship could also be extended functionally. Based on our findings, ventral BA44 is associated with verbal working memory processes, i.e. prefrontal functions. On the other hand, dorsal BA44 is sensitive to length effects independent of working memory demands. As we've mentioned in chapter 4, length effects generally reflect that a region is part of the phonological/phonetic analysis system, which suggests that it is involved in some of the processes that will lead to the generation of an articulatory plan.

The exact role that dorsal BA44 plays in this process cannot be fully specified based on our results. The fact that dorsal BA44 is sensitive to length effects, independent of task delay, suggests that it could be involved in processes related to sequencing and syllabification, as described by Friederici (2002). Such an account would also be in partial agreement with Indefrey and Levelt (2000; 2004) and the hypothesis that the LIFG is involved in phonological processing and in particular syllabification. With respect to the role of the ventral part of the LIFG, our findings suggest that it is probably related to verbal working memory. By further showing that under conditions of task delay this region is also sensitive to differences in sub-lexical frequency, we have also extended previous results on the role of the ventral LIFG. So

far, previous research has only showed that the ventral LIFG is sensitive to lexicality, e.g. non-words vs. words (Bokde et al. 2001; Chein and Fiez 2001; Chein et al. 2002). However, it is possible that these results were biased by a potential difference in sub-lexical frequency between words and non-words, with non-words possibly consisting of lower frequency components than real words. In this sense it is not lexicality that is a modulator of the activity of ventral BA44 during verbal working memory tasks, but rather frequency of occurrence of the target's components.

As is the case with the generation of articulatory codes for language production, low sub-lexical frequency targets might require more resources and effort during sub-vocal rehearsal, since their articulatory codes would need to be compiled on-line. Based on the fact that there is a clear relationship between language production and sub-vocal rehearsal (Caplan and Waters 1995; Baddeley 2003), it seems that the effort and resource demand that takes place at the stage of compiling the articulatory codes could also affect the stage of sub-vocal rehearsal in a top-down manner. Even though a PPI analysis did not reveal any significant changes in connectivity between the LIFG and the precentral gyrus, this does not necessarily mean that the two regions are not functionally interacting with one another. It could also be pointing to the fact that the connectivity of the two regions is not dependent on the type of information that they are exchanging and that it could be task dependent, e.g. the regions would cooperate under conditions of verbal working memory only. In this case, we would not expect to see stimulus related changes in the PPI results, but only task-related changes. Such questions could possibly be addressed in future experiments and analyses.

7.4 Future Work

In this thesis, we have presented the work that we have conducted to examine the neuroanatomical substrates of phonological and phonetic encoding and in particular the role of the LIFG in these processes. Beyond this work, there are still many issues and questions that need to be addressed. In this final section, we would like to summarize some of the directions that we hope to explore in the future and some of the questions that we feel should be addressed in future experiments.

For one thing, it would be interesting to examine the single-subject activation patterns and see whether we could observe the same dorsal-ventral segregation pattern as we see in the group results at a single-subject level. Even though we were able to replicate the functional segregation of the LIFG in the group data from two different population samples, a single-subject analysis could reveal more precise information about the anatomical features of the activation patterns. The high spatial resolution data could potentially be used for such an analysis, in particular the un-smoothed data, though one would need to be very cautious about interpreting these results. Unless all subjects show similar activation patterns, the same subjects would need to be re-scanned to verify the validity of any single-subject assumptions.

Furthermore, in all the analyses employed we used an activation-based approach. In doing so, we made the assumption that any differences between the processing of contrasting conditions, e.g. low and high phonotactic probability pseudowords, would show up as differences in the magnitude and extent of the activation of the region. However, such an approach effectively ignores regions that are involved in the processing of the two conditions, i.e. those encoding information relevant to the processing of the two conditions, but which do not show any difference in the magnitude of the activation induced by the condition.

In order to identify such regions, an information-based analysis should be applied as described in chapter 6. In our case, such an approach could further clarify whether the ventral part of the LIFG is only involved in verbal working memory or whether it is generally engaged in the processing of verbal stimuli. If it can be shown that during prompt response trials the ventral LIFG contains information about both types of stimuli (low and high sub-lexical frequency pseudowords), then this would mean that its function is not strictly related to verbal working memory processing, but for some reason the differences between the two conditions are emphasized under verbal working memory demands.

Finally, future experiments would also need to be performed to address more refined questions about the type of processes that are taking place within the two LIFG subregions. In the studies that were presented as part of this thesis, we used the effects of length and sub-lexical frequency to test for regions involved in phonological and phonetic encoding. With respect to sub-lexical frequency, we did not differentiate between the frequency of syllables and biphones and the frequency of phonemes. Based on some of the proposed models on phonetic encoding (Indefrey and Levelt 2000; 2004; Guenther et al. 2006), one would not expect that the sub-lexical frequency of the phonemes would cause a difference in the magnitude of the activation of the regions involved in phonetic encoding. Differences in phonetic encoding arise as a result of compiling or retrieving different articulatory codes for syllables or any other complex articulatory unit. Individual phonemes would be pre-compiled, irrespective of whether they have low or high sub-lexical frequency, as suggested in the DIVA computational model of language production presented by Guenther et al. (2006). However, what is not clear is what the situation is for verbal working memory related processes. In our case we observed differences in activation between low and high sub-lexical frequency pseudowords specific to a verbal working memory task. If it is the case that verbal working memory processes, such as sub-vocal rehearsal, recruit mechanisms related to phonetic encoding, then

we would not expect that the sub-lexical frequency of the phonemes would be driving the differences that we observed in ventral LIFG. However, this is a hypothesis that would need to be tested.

To this extent we have shown that the effects observed in ventral BA44 for contrasts such as words vs. non-words could be attributed to differences in sub-lexical frequency between the two categories (words and non-words). At the same time it would also be necessary to examine the source of the sub-lexical frequency effect. Is it because of the frequency of the phonemes or of the syllables? If we can identify the syllables as the source of this difference, then it would mean that the same process of compiling articulatory codes for language production also affects verbal working memory processes and sub-vocal rehearsal. In the opposite case, it would suggest that sub-vocal rehearsal employs a different mechanism to generate articulatory codes than specified for language production. To answer this question, future experiments using more tightly controlled stimulus pairs would need to be conducted.

Chapter 8 Conclusion

In this thesis, we have presented work that we have conducted on the study of the phonological/phonetic system, its neuroanatomical substrates and the role of the left inferior frontal gyrus (LIFG). Even though there is much more work to be done before we can specify the exact details of the system and the role of the regions involved, we were able to disambiguate some of the contrasting points within proposed neuroanatomical models of language processing. In particular, we identified the precentral gyrus as a key region in the process of phonetic encoding and the compilation of articulatory codes as it has been proposed in the models of Hickok and Poeppel (2000; 2004; 2007). We propose that this area is the storage site of articulatory codes in agreement with theories on the existence of a mental syllabary or speech sound map. We further showed that BA44, the posterior part of Broca's area, is functionally segregated. The dorsal part of BA44, only showed an effect of pseudoword length, suggesting that it has a role in phonological processing as has been claimed by Indefrey and Levelt (2000; 2004). The ventral part on the other hand showed both an effect of length and sub-lexical frequency. Preliminary evidence also suggest that there may be an interaction between sub-lexical frequency and response delay, which would be in agreement with a role of the LIFG in verbal working memory and covert rehearsal as it has been proposed by Baddeley (2003). This will need to be further verified in future research. This evidence brings further support to the claim that the LIFG, and even more so BA44, is not a functionally homogeneous region and current neuroanatomical models of linguistic and non-linguistic processing should be revised to take into account the functional segregation of the LIFG and BA44 in particular. Generating more detailed models on the relationship between function and structure within the LIFG would be a step forward both in disambiguating its role and in understanding linguistic processing in the brain.

Bibliography

Amunts K, Schleicher A, Burgel U, Mohlberg H, Uylings HB, Zilles K. 1999. Broca's region revisited: Cytoarchitecture and intersubject variability. Journal of Comparative Neurology. 412(2):319-341.

Amunts K, Weiss PH, Mohlberg H, Pieperhoff P, Eickhoff SB, Gurd JM, Marshall JC, Shah NJ, Fink GR, Zilles K. 2004. Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space-the roles of Brodmann areas 44 and 45. Neuroimage. 22(1):42-56.

Amunts K, Zilles K. 2006. A Multimodal Analysis of Structure and Function in Broca's Region. In: Grodzinsky Y, Amunts K, editors. Broca's Region. Oxford University Press; pp. 17-30.

Anwander A, Tittgemeyer M, Cramon DV, Friederici AD, Knösche T. 2007. Connectivity-Based Parcellation of Broca's Area. Cerebral Cortex. 17(4):816-825.

Arbib M. 2003. Language evolution: The mirror system hypothesis. In: The handbook of brain theory and neural networks. MIT Press; pp. 606-611.

Bachoud-Lévi A, Dupoux E, Cohen L, Mehler J. 1998. Where Is the Length Effect? A Cross-Linguistic Study of Speech Production. Journal of Memory and Language. 39(3):331-346.

Baddeley A. 1966. The influence of acoustic and semantic similarity on long-term memory for word sequences. Quarterly Journal of Experimental Psychology. 18(4):302-309.

Baddeley A. 1992. Working memory. Science. 255(5044):556-559.

Baddeley A. 2003. Working memory and language: an overview. Journal of Communication Disorders. 36(3):189-208.

Bammer R, Keeling SL, Augustin M, Pruessmann KP, Wolf R, Stollberger R, Hartung HP, Fazekas F. 2001. Improved diffusion-weighted single-shot echo-planar imaging (EPI) in stroke using sensitivity encoding (SENSE). Magnetic Resonance in Medicine. 46(3):548-554.

Barch DM, Sabb FW, Carter CS, Braver TS, Noll DC, Cohen JD. 1999. Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. Neuroimage. 10(6):642-657.

Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A. 2004. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nature Neuroscience. 7(11):1190-1192.

Belin P, Zatorre RJ, Hoge R, Evans AC, Pike B. 1999. Event-related fMRI of the auditory cortex. Neuroimage. 10(4):417-429.

Benardete EA, Victor JD. 1994. An extension of the m-sequence technique for the analysis of multi-input nonlinear systems. In: Marmarelis VZ, editor. Vol. 3. Plenum Press, New York; pp. 87-110.

Binkofski F, Buccino G. 2004. Motor functions of Broca's region. Brain and Language. 89(2):362-9.

Birn RM, Bandettini PA, Cox RW, Shaker R. 1999. Event-related fMRI of tasks involving brief motion. Human Brain Mapping. 7(2):106-114.

Birn RM, Cox RW, Bandettini PA. 2004. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. Neuroimage. 23(3):1046-1058.

Bohland JW, Guenther FH. 2006. An fMRI investigation of syllable sequence production. Neuroimage. 32(2):821-841.

Bokde AL, Tagamets MA, Friedman RB, Horwitz B. 2001. Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. Neuron. 30(2):609-617.

Brett M, Anton J, Valabregue R, Poline J. 2002. Region of interest analysis using an SPM toolbox. In: Proceedings of the 8th International Conference on Functional Mapping of the Human Brain in Sendai, Japan. Available on CD-ROM in Neuroimage. Vol. 16.

Broca P. 1861. Remarques sur le siege de la faculte de langage articule, suivis d'une observation d'aphemie (perte de la parole). Bulletins de la Societe Anatomique. 6:330-357.

Browman CP, Goldstein L. 1988. Some notes on syllable structure in articulatory phonology. Phonetica. 45(2-4):140-155.

Buccino G, Binkofski F, Fink GR, Fadiga L, Fogassi L, Gallese V, Seitz RJ, Zilles K, Rizzolatti G, Freund HJ. 2001. Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. European Journal of Neuroscience. 13(2):400-404.

Buccino G, Binkofski F, Riggio L. 2004. The mirror neuron system and action recognition. Brain and Language. 89(2):370-376.

Buracas GT, Boynton GM. 2002. Efficient design of event-related fMRI experiments using M-sequences. Neuroimage. 16(3 Pt 1):801-813.

Burton MW, Small SL, Blumstein SE. 2000. The role of segmentation in phonological processing: An fMRI investigation. Journal of Cognitive Neuroscience. 12(4):679-90.

Buxton RB. 2002. Introduction to functional magnetic resonance imaging. Cambridge University Press

Caplan D. 2001. Functional Neuroimaging Studies of Syntactic Processing. Journal of Psycholinguistic Research. 30(3):297-320.

Caplan D, Waters GS. 1995. On the nature of the phonological output planning processes involved in verbal rehearsal: evidence from aphasia. Brain and Language. 48(2):191-220.

Cavanna AE, Trimble MR. 2006. The precuneus: a review of its functional anatomy and behavioural correlates. Brain. 129(Pt 3):564-583.

Chein JM, Fiez JA. 2001. Dissociation of verbal working memory system components using a delayed serial recall task. Cerebral Cortex. 11(11):1003-1014.

Chein JM, Fissell K, Jacobs S, Fiez JA. 2002. Functional heterogeneity within Broca's area during verbal working memory. Physiology and Behavior. 77(4-5):635-639.

Cheng K, Waggoner RA, Tanaka K. 2001. Human ocular dominance columns as revealed by high-field functional magnetic resonance imaging. Neuron. 32(2):359-374.

Chen W, Ogawa S. 1999. Principle of BOLD-functional MRI. In: Moonen CTW, Bandettini PA, editors. Functional MRI. Springer; pp. 103-114.

Cholin J, Levelt WJ, Schiller NO. 2006. Effects of syllable frequency in speech production. Cognition. 99:205-235.

Dale AM, Buckner RL. 1997. Selective averaging of rapidly presented individual trials using fMRI. Human Brain Mapping. 5:329-340.

Demonet JF, Fiez JA, Paulesu E, Petersen SE, Zatorre RJ. 1996. PET Studies of Phonological Processing: A Critical Reply to Poeppel. Brain and Language. 55(3):352-379.

Devlin JT, Matthews PM, Rushworth MFS. 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.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage. 25(4):1325-1335.

Fadiga L, Craighero L, Buccino G, Rizzolatti G. 2002. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. European Journal of Neuroscience. 15(2):399-402.

Fiez JA, Balota DA, Raichle ME, Petersen SE. 1999. Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. Neuron. 24(1):205-218.

Fox PT, Raichle ME, Mintun MA, Dence C. 1988. Nonoxidative glucose consumption during focal physiologic neural activity. Science. 241:462-464.

Friederici AD. 2002. Towards a neural basis of auditory sentence processing. Trends in Cognitive Science. 6(2):78-84.

Friederici AD. 2006. Broca's area and the ventral premotor cortex in language: functional differentiation and specificity. Cortex. 42(4):472-475.

Frisch SA, Large NR, Pisoni DB. 2000. Perception of Wordlikeness: Effects of Segment Probability and Length on the processing of nonwords. Journal of Memory and Language. 42:481-496.

Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. 1997. Psychophysiological and modulatory interactions in neuroimaging. Neuroimage. 6(3):218-229.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. 1998. Event-related fMRI: characterizing differential responses. Neuroimage. 7(1):30-40.

Friston KJ, Mechelli A, Turner R, Price CJ. 2000. Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. Neuroimage. 12(4):466-477.

Friston KJ. 1994. Functional and effective connectivity in neuroimaging: A synthesis. Human Brain Mapping. 2(1-2):56-78.

Galantucci B, Fowler CA, Turvey MT. 2006. The motor theory of speech perception reviewed. Psychonomic Bulletin and Review. 13(3):361-377.

Gallese V, Fadiga L, Fogassi L, Rizzolatti G. 1996. Action recognition in the premotor cortex. Brain. 119(Pt 2):593-609.

Gelfand JR, Bookheimer SY. 2003. Dissociating neural mechanisms of temporal sequencing and processing phonemes. Neuron. 38(5):831-842.

Gibson JJ. 1979. The ecological approach to visual perception. Hillsdale, NJ: Lawrence Erlbaum.

Gitelman DR, Penny WD, Ashburner J, Friston KJ. 2003. Modeling regional and psychophysiologic interactions in fMRI: the importance of hemodynamic deconvolution. Neuroimage. 19(1):200-207.

Goldrick M, Rapp B. 2007. Lexical and post-lexical phonological representations in spoken production. Cognition. 102(2):219-260.

Graves WW, Grabowski TJ, Mehta S, Gordon JK. 2007. A neural signature of phonological access: distinguishing the effects of word frequency from familiarity and length in overt picture naming. Journal of Cognitive Neuroscience. 19(4):617-631.

Grezes J, Armony JL, Rowe J, Passingham RE. 2003. Activations related to "mirror" and "canonical" neurones in the human brain: An fMRI study. Neuroimage. 18(4):928-37.

Guenther FH, Ghosh SS, Tourville JA. 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain and Language. 96(3):280-301.

Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. 1999. "Sparse" temporal sampling in auditory fMRI. Human Brain Mapping. 7(3):213-223.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science. 293(5539):2425-2430.

Hayasaka S, Nichols TE. 2003. Validating cluster size inference: random field and permutation methods. Neuroimage. 20(4):2343-2356.

Heim S, Opitz B, Friederici AD. 2003a. Distributed cortical networks for syntax processing: Broca's area as the common denominator. Brain and Language. 85(3):402-408.

Heim S, Opitz B, Muller K, Friederici AD. 2003b. Phonological processing during language production: fMRI evidence for a shared production-comprehension network. Cognitive Brain Research. 16(2):285-296.

Henson R. 2003. Analysis of fMRI Time Series Linear Time-Invariant Models, Event-Related fMRI and Optimal Experimental Design. In: Frackowiak RS, Friston KJ, Frith CD, Dolan RJ, Price CJ, Zeki S, Ashburner JT, Penny WD, editors. Human Brain Function. Vol. 2. 2 ed. Academic Press; pp. 793-822. Available from: http://www.sciencedirect.com/science/book/9780122648410

Hickok G, Buchsbaum B, Humphries C, Muftuler T. 2003. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. Journal of Cognitive Neuroscience. 15(5):673-682.

Hickok G, Poeppel D. 2000. Towards a functional neuroanatomy of speech perception. Trends in Cognitive Science. 4(4):131-138.

Hickok G, Poeppel D. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition. 92(1-2):67-99.

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. Nature Reviews Neuroscience. 8(5):393-402.

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

Iacoboni M. 2005. Understanding others: Imitation, language, empathy. In: Perspectives on imitation: from cognitive neuroscience to social science. Vol. 1. Cambridge, MA: MIT Press.

Iacoboni M, Woods RP, Brass M, Bekkering H, Mazziotta JC, Rizzolatti G. 1999. Cortical mechanisms of human imitation. Science. 286(5449):2526-2528.

Indefrey P, Levelt WJ. 2000. The new cognitive neurosciences. In: Gazzaniga M, editor. Cambridge, MA: MIT Press; pp. 845-865.

Indefrey P, Levelt WJ. 2004. The spatial and temporal signatures of word production components. Cognition. 92(1-2):101-144.

International Phonetic Association. 1999. Handbook of the International Phonetic Association: A guide to the use of the international phonetic alphabet. Cambridge University Press

Jusczyk P, Luce PA, Charles-Luce J. 1994. Infants' sensitivity to phonotactic patterns in the native language. Journal of Memory and Language. 33:630-645.

Kellman P, van Gelderen P, de Zwart JA, Duyn JH. 2003. Method for functional MRI mapping of nonlinear response. Neuroimage. 19(1):190-199.

Kriegeskorte N, Bandettini P. 2007. Analyzing for information, not activation, to exploit high-resolution fMRI. Neuroimage. 38(4):649-662.

Kriegeskorte N, Goebel R, Bandettini P. 2006. Information-based functional brain mapping. Proceedings of The National Academy Of Sciences Of The United States Of America. 103(10):3863-3868.

Lai S, Glover GH, Haacke EM. 1999. Spatial selectivity of BOLD contrast: effects in and around draining veins. In: Moonen CTW, Bandettini PA, editors. Functional MRI. Springer; pp. 221-231.

Levelt WJ. 1999. Models of word production. Trends in Cognitive Science. 3(6):223-232.

Levelt WJ, Roelofs A, Meyer AS. 1999. A theory of lexical access in speech production. Behavioral and Brain Sciences. 22(1):1-38; discussion 38-75.

Levelt WJ, Wheeldon L. 1994. Do speakers have access to a mental syllabary? Cognition. 50(1-3):239-269.

Liberman AM, Mattingly IG. 1985. The motor theory of speech perception revised. Cognition. 21(1):1-36.

Liu TT, Frank LR, Wong EC, Buxton RB. 2001. Detection power, estimation efficiency, and predictability in event-related fMRI. Neuroimage. 13(4):759-773.

Magistretti PJ, Pellerin L. 1999. Regulation of Cerebral Energy Metabolism. In: Moonen CTW, Bandettini PA, editors. Functional MRI. Springer; pp. 25-35.

Makuuchi M. 2005. Is Broca's area crucial for imitation? Cerebral Cortex. 15(5):563-570.

Mazaika P, Whitfield-Gabrieli S, Reiss A. 2007. Artifact Repair for fMRI Data from High Motion Clinical Subjects. In: Human Brain Mapping Conference.

Menon RS, Ogawa S, Strupp JP, Ugurbil K. 1997. Ocular dominance in human V1 demonstrated by functional magnetic resonance imaging. Journal of Neurophysiology. 77(5):2780-2787.

Mesulam MM, Mufson EJ. 1982. Insula of the old world monkey. I. Architectonics in the insulo-orbito-temporal component of the paralimbic brain. Journal of Comparative Neurology. 212(1):1-22.

Mikl M, Marecek R, Hlustík P, Pavlicová M, Drastich A, Chlebus P, Brázdil M, Krupa P. 2008. Effects of spatial smoothing on fMRI group inferences. Magnetic Resonance Imaging. 26(4):490-503.

Milner AD, Goodale MA. 1995. The visual brain in action. Oxford University Press

Molnar-Szakacs I, Iacoboni M, Koski L, Mazziotta JC. 2005. Functional segregation within pars opercularis of the inferior frontal gyrus: evidence from fMRI studies of imitation and action observation. Cerebral Cortex. 15(7):986-994.

Murata A, Fadiga L, Fogassi L, Gallese V, Raos V, Rizzolatti G. 1997. Object representation in the ventral premotor cortex (area F5) of the monkey. Journal of Neurophysiology. 78(4):2226-2230.

Nencka AS, Rowe DB. 2007. Reducing the unwanted draining vein BOLD contribution in fMRI with statistical post-processing methods. Neuroimage. 37(1):177-188.

Newman SD, Just MA, Keller TA, Roth J, Carpenter PA. 2003. Differential effects of syntactic and semantic processing on the subregions of Broca's area. Cognitive Brain Research. 16(2):297-307.

Nishitani N, Schurmann M, Amunts K, Hari R. 2005. Broca's Region: From Action to Language. Physiology. 20:60-69.

Oakes TR, Johnstone T, Walsh KSO, Greischar LL, Alexander AL, Fox AS, Davidson RJ. 2005. Comparison of fMRI motion correction software tools. Neuroimage. 28(3):529-543.

Ohbayashi M, Ohki K, Miyashita Y. 2003. Conversion of Working Memory to Motor Sequence in the Monkey Premotor Cortex. Science. 301:233-236.

Okada K, Hickok G. 2006a. Left posterior auditory-related cortices participate both in speech perception and speech production: Neural overlap revealed by fMRI. Brain and Language. 98(1):112-117.

Okada K, Hickok G. 2006b. Identification of lexical-phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. Neuroreport. 17(12):1293-1296.

Oztop E, Arbib MA. 2002. Schema design and implementation of the grasp-related mirror neuron system. Biological Cybernetics. 87(2):116-140.

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. 1992. Understanding motor events: A neurophysiological study. Experimental Brain Research. 91(1):176-180.

Petrides M. 2006. Broca's area in the human and the nonhuman primate brain. In: Grodzinsky Y, Amunts K, editors. Broca's Region. Oxford university press; pp. 31-46.

Petrides M, Cadoret G, Mackey S. 2005. Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. Nature. 435(7046):1235-1238.

Pisoni DB, Nusbaum HC, Luce PA, Slowiaczek LM. 1985. Speech perception, word recognition and the structure of the lexicon. Speech Communication. 4(1-3):75-95.

Poldrack RA, Temple E, Protopapas A, Nagarajan S, Tallal P, Merzenich M, Gabrieli JD. 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 RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD. 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage. 10(1):15-35.

Pulvermüller F, Huss M, Kherif F, Martin FMDP, Hauk O, Shtyrov Y. 2006. Motor cortex maps articulatory features of speech sounds. Proceedings of The National Academy Of Sciences Of The United States Of America. 103(20):7865-7870.

Riecker A, Mathiak K, Wildgruber D, Erb M, Hertrich I, Grodd W, Ackermann H. 2005. fMRI reveals two distinct cerebral networks subserving speech motor control. Neurology. 64(4):700-706.

Rizzolatti G, Arbib MA. 1998. Language within our grasp. Trends in Neurosciences. 21(5):188-194.

Rizzolatti G, Craighero L. 2004. The mirror-neuron system. Annual Reviews - Neuroscience. 27:169-192.

Rizzolatti G, Fogassi L, Gallese V. 2002. Motor and cognitive functions of the ventral premotor cortex. Current Opinion in Neurobiology. 12(2):149-154.

Rizzolatti G, Fogassi L, Galles V. 1997. Parietal cortex: from sight to action. Current Opinion In Neurobiology. 7(4):562-567.

Sakata H, Taira M, Murata A, Mine S. 1995. Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey. Cerebral Cortex. 5(5):429-438.

Savoy RL, Bandettini PA, O'Craven KM, Kwong KK, Davis TL, Baker JR, Weiskoff RM, Rosen BR. 1995. Pushing the temporal resolution of fMRI: Studies of very brief visual stimuli, onset variability and asynchrony, and stimulus-correlated changes in noise. In: Proceedings of the Society of Magnetic Resonance Third Scientific Meeting and Exhibition. Vol. 2.; p. 450.

Shuster LI, Lemieux SK. 2005. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. Brain and Language. 93(1):20-31.

Springer CJ, Patlak CS, Palyka I, Huang W. 1999. Principles of susceptibility contrast-based functional MRI: the sign of the functional MRI response. In: Moonen CTW, Bandettini PA, editors. Functional MRI. Springer; pp. 91-102.

Strand F, Forssberg H, Klingberg T, Norrelgen F. 2008. Phonological working memory with auditory presentation of pseudo-words — an event related fMRI Study. Brain Research. 1212:48-54.

Tagamets MA, Novick JM, Chalmers ML, Friedman RB. 2000. A parametric approach to orthographic processing in the brain: An fMRI study. Journal of Cognitive Neuroscience. 12(2):281-297.

Vallar G, Papagno C. 2002. Neuropsychological impairments of verbal short-term memory. In: Baddeley, A. D., Kopelman, M. D. , Wilson, B. A., editors. Handbook of memory disorders. 2 ed. Chichester: Wiley; pp. 249-270.

Villringer A. 1999. Physiological changes during brain activation. In: Moonen CTW, Bandettini PA, editors. Functional MRI. Springer; pp. 3-14.

Vitevitch MS, Luce PA. 1998. When Words Compete: Levels of processing in perception of spoken words. Psychological Science. 9(4):325-329.

Vitevitch MS, Luce PA. 2004. A web-based interface to calculate phonotactic probability for words and nonwords in English. Behavior Research Methods, Instruments, & Computers. 36(3):481-487.

Vitevitch MS, Luce PA. 2005. Increases in phonotactic probability facilitate spoken nonword repetition. Journal of Memory and Language. 52(2):193-204.

Vitevitch MS, Luce PA, Charles-Luce J, Kemmerer D. 1997. Phonotactics and syllable stress: implications for the processing of spoken nonsense words. Language and Speech. 40(Pt 1):47-62.

Vitevitch MS, Luce PA, Pisoni DB, Auer ET. 1999. Phonotactics, neighborhood activation, and lexical access for spoken words. Brain and Language. 68(1-2):306-311.

Wilson M. 2001. The case for sensorimotor coding in working memory. Psychonomic Bulletin and Review. 8(1):44-57.

Zaehle T, Schmidt CF, Meyer M, Baumann S, Baltes C, Boesiger P, Jancke L. 2007. Comparison of "silent" clustered and sparse temporal fMRI acquisitions in tonal and speech perception tasks. Neuroimage. 37(4):1195-1204.

Zatorre RJ, Evans AC, Meyer E, Gjedde A. 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science. 256(5058):846-849.

Zatorre RJ, Meyer E, Gjedde A, Evans AC. 1996. PET studies of phonetic processing of speech: Review, replication, and reanalysis. Cerebral Cortex. 6(1):21-30.

de Zwart JA, van Gelderen P, Kellman P, Duyn JH. 2002. Application of sensitivity-encoded echo-planar imaging for blood oxygen level-dependent functional brain imaging. Magnetic Resonance in Medicine. 48(6):1011-1020.

de Zwart JA, Ledden PJ, van Gelderen P, Bodurka J, Chu R, Duyn JH. 2004. Signal-to-noise ratio and parallel imaging performance of a 16-channel receive-only brain coil array at 3.0 Tesla. Magnetic Resonance in Medicine. 51(1):22-26.

Appendices

A Stimuli

Two-syllable (CV.CVC)		Four-syllable (CV.CV.CV.CVC)	
High	Low	High	Low
sɪsəs	yɔyəʃ	hasɜˠsasɛd	gɔɪyaθoyʌtʃ
sɪsəl	gɔɪtʃɜˠz	matæsasæm	yɔɪwɛʃoθʌtʃ
kɪnɛn	gɔɪʃɛg	hɛtəˠsɛsædʒ	zuwaʃoðaɪp
sʌsət	gɔɪʃɜˠz	hɛsætɛsɜˠl	ycwaʃoθʌdʒ
kɛrɪt	gɔɪʃaɪm	mɪrætɛsʌg	gɔɪyɛθowʌtʃ
kɛnɛt	gɔɪwʌdʒ	hɪrætasaɪb	zɔɪyæʃoθɛθ
kɪsɜˠn	gɔɪyʌdʒ	hatɜˠsakæn	gɔɪfaθoðʌtʃ
hɪnin	gɔɪyʌz	hɛtətɛsɜˠg	vuwaθəˠðʌg
kɪsaɪs	gɔɪyɜˠp	hasæsɛkɪs	yuwɛʃoθɜˠθ
kɛlæn	gɔɪyɜˠz	hɪsætadəp	gɔɪyaθoyʌdʒ
sɪsɜˠl	gutʃʌtʃ	mɪtɪsɛsæv	yɔɪwaʃəˠtzɜˠdʒ
halən	gutʃɛð	hɛsəˠsɛsɜˠk	vcwæθəˠdʒɜˠθ
mɛnɛn	gutʃɛθ	matɪtaləp	gɔɪyɛθowʌdʒ
mɛrɪn	guðʌv	sitəsɛsɛp	vuwaθəˠdʒaɪð
kɛdɪt	guðoɪl	mɛtɪtɛsaɪl	yuyəˠʃəˠtʃɜˠg
hɪtəl	guðaɪb	kɪsætɛmaɪd	yuwaʃəˠtʃɜˠθ
kɪkɜˠn	guðaɪp	mɛsəˠtɛsaɪp	yoyaʃəˠʒoɪl
kɪsaɪl	gudʒɛg	sitɪtɛsɛb	vuwɛθəˠyɜˠp
sɪsɛb	gudʒɜˠθ	kɛtɪsɛsaɪm	guwaθoðaɪð
sɪsɜˠd	vctʃaɪb	kɪtɪsakɛd	guyɛθoðʌdʒ
kɛnɛm	vcʃʌdʒ	mɛsætadɪk	vuyaθəˠðʌd
sɪvaɪt	vcʃɜˠtʃ	kɪtɪsɛsɜˠp	zuwɛʃoðaɪb
sɪsɜˠm	vcʃaɪð	hɛrətɛsɜˠd	gowɛθoyɜˠtʃ
hakət	vcθʌdʒ	mɛtætapaɪt	guyaθoðʌdʒ
mɛləm	vcθʌz	kɪtɜˠsaləl	gowɛθoyɜˠz
sʌsəv	vuðaɪð	hɪtɪsasiv	yuyɛʃoθʌz
sɪsʌg	vcθɜˠθ	mɪsɛtɛsʌʃ	vuyɛθəˠðaɪm
makɛd	vcθuʃ	kɪtəˠsɛsɜˠm	gofaθoyəˠg
hɪmaɪd	vcʒoɪl	kɪsəˠsakɛb	vuyɛθəˠðoɪl
masaɪp	vuðʌdʒ	kɛtɪtɛsaɪv	goɪfɛθoðʌv
masɜˠp	vuðʌz	harəsɛkɪk	gowaθoyʌz
sʌsəz	vudʒɛð	sɪtɜˠtanæm	yuwɛʃəˠtʃʌtʃ
masəd	vudʒɛθ	sasəˠtarɪg	zoɪwɛʃəˠʃaɪð
kɪsaɪm	vuθɛθ	kɪsæsasig	zoɪfæʃəˠʃɜˠtʃ
sʌsɜˠt	vuyɜˠtʃ	hatɪsadɪt	yuyɛʃəˠtʃaɪð
matɪn	vuyɜˠg	mitəˠtɛsɜˠz	gofɛtoyɜˠθ
mɛris	vuyɜˠθ	mɪsəˠsɛsɜˠs	vufaθəˠtʃʌdʒ
mɛləp	ycðʌd	matəˠtanəp	vuyaθəˠdʒɛdʒ

kɛdəp	ycðʌg	mɛrætalæn	zuyɛʃoʃaɪp
sasɪv	ycðʌʃ	mɪtɚsɛsəz	yoɪyæʃɚtʃɛdʒ
haləl	ycðið	hɛtɪsɛsæt	yoɪfɛʃoθɛð
kɛnæp	ycwʌʃ	hatæsɛtɜ˞t	zoɪwaʃɚʃɛð
kɛrɪg	ycyɪtʃ	hɪsɚsasɪs	voɪyɛθɚdʒɛg
kɪsaɪb	ycyidʒ	hɪtɜ˞sakət	zufaʃoʃʌtʃ
sɪsɜ˞g	ycyɪθ	mɛtɚtɛdis	vcwɛθɚwʌʃ
hɪtɜ˞t	yctʃɛdʒ	satətarɪz	zoɪyɛʃɚʃɛθ
kɪsɜ˞z	yctʃɜ˞dʒ	kɪtatɛsəd	yoɪyɛʃɚtʃʌʃ
kɪkɛb	yoɪðʌʃ	kɛsɚtɛsaɪk	zufæʃotʃaɪb
hɪtɛs	yoɪʃʌtʃ	hɪtætakɪt	vuwɛθɚðʌʃ
kɛʃəm	yoɪʃɛð	masɜ˞talən	zoɪfɛʃɚʃɜ˞z
masɛs	yoɪʃɛθ	satæsarɪt	yoɪfaʃɚtʃɛg
hanɛs	yoɪʃɜ˞θ	kɛtɚtɛsəl	zufɛʃoʃʌdʒ
hakɪt	yoɪʃaɪb	kɪtætanɛs	zuyæʃoyidʒ
mɛdɪs	yoɪʃaɪp	kɛsəsɛsaɪs	vofaθɚyʌʃ
sʌsaɪk	zoɪtʃʌdʒ	sɪsæsaris	zuyɛʃoʃaɪm
sʌsɜ˞k	zoɪtʃʌz	sɪsɚtaləm	guwɛθodʒɛð
mɛdæt	zoɪtʃɜ˞g	sɪtɚtɛtəl	goɪwaθoyɪð
hɪpaɪt	zoɪtʃɜ˞θ	sɪsɜ˞tɛkɜ˞n	zuwaʃoʃɛg
mɛrɪz	zoɪtʃaɪð	kɪrætatɛs	yufaʃɚtʃɛð
medæn	zoɪðʌtʃ	sɪtɪtadæt	vufɛθɚtʃʌz
mavin	zoɪwʌtʃ	sɪsætavɪn	gufɛθodʒɛθ
hakɪs	zoɪyʌtʃ	masəsanɛn	ycfɛʃɚtʃɜ˞z
masɛd	zoɪyið	kɪsɚsɛsəs	zufɛʃoʃaɪb
hɪdik	zutʃʌʃ	sɪræsanæt	yufɛʃɚtʃɛθ
hakæn	zutʃɛg	kɛrxsɛsɜ˞n	ycfaʃoθuʃ
sʌsaɪv	zuðɑʃ	kɪrətɛvaɪt	vuyɛθɚyɪθ
masɛp	zuθʌtʃ	satɪtadən	vufaθɚðɪð
sʌsɜ˞z	zuθɛð	sɪrxsɛsɜ˞t	gufɛθoðuʃ
hakɪk	zuʒaʃ	sasəsarin	vofɛθɚyɪtʃ
hasɪg	zuðaɪm	marəsatɪn	zuwɛʃɚʒʌʃ
hɪsʌʃ	zudʒɛdʒ	særəsanæn	vufɛθɚðʌʃ
hɪsʌdʒ	zudʒaɪð	sɪtəsanɪn	zuyæʃoʃɜ˞θ

Note: The stimuli are phonetically transcribed based on the International Phonetic Association (1999).

B Division of Labour

The author, Marina Papoutsi, designed the fMRI paradigm and generated the stimuli for all experiments performed, as well as acquired and analysed all experimental data and wrote the submitted paper. Dr. Martijn Jansma (NIMH, NIH, Bethesda, MD) was involved in designing the fMRI paradigm by providing the m-sequence files and help in using the m-sequences. Dr. Jacco De Zwart (NINDS, NIH, Bethesda, MD) was involved in designing the fMRI scanning protocol by providing the EPI sequence and support during the image acquisition. Drs. Barry Horwitz, Martin Pickering and James Bednar had the role of supervising the experiments and had overview of all stages of the experiments and the resulting scientific publications.

C Publications Arising from this Thesis

M. Papoutsi, J.A. de Zwart, J.M. Jansma, M. Pickering, J. A. Bednar and B. Horwitz, "The Processing of low frequency pseudowords by Broca's area", *Organisation for Human Brain Mapping, 13th Annual Meeting, Chicago, USA*, 2007.

M. Papoutsi, J.A. de Zwart, J.M. Jansma, M. Pickering, J. A. Bednar and B. Horwitz, "From phonemes to articulatory codes: an fMRI study of the role of Broca's area in speech production", *Cerebral Cortex 2009; doi: 10.1093/cercor/bhn239.*

Cerebral Cortex Advance Access published January 29, 2009

Cerebral Cortex
doi:10.1093/cercor/bhn239

From Phonemes to Articulatory Codes: An fMRI Study of the Role of Broca's Area in Speech Production

Marina Papoutsi[1,2], Jacco A. de Zwart[3], J. Martijn Jansma[4], Martin J. Pickering[5], James A. Bednar[1] and Barry Horwitz[2]

[1]Institute for Adaptive and Neural Computation, University of Edinburgh, UK, [2]Brain Imaging Modeling Section, Voice, Speech and Language Branch, National Institute on Deafness and Other Communication Disorders, National Institutes of Health, Bethesda, MD, USA, [3]Advanced MRI Section, Laboratory of Functional and Molecular Imaging, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA, [4]Neuroimaging Section, Mood and Anxiety Disorders Program, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA and [5]Department of Psychology, University of Edinburgh, UK

We used event-related functional magnetic resonance imaging to investigate the neuroanatomical substrates of phonetic encoding and the generation of articulatory codes from phonological representations. Our focus was on the role of the left inferior frontal gyrus (LIFG) and in particular whether the LIFG plays a role in sublexical phonological processing such as syllabification or whether it is directly involved in phonetic encoding and the generation of articulatory codes. To answer this question, we contrasted the brain activation patterns elicited by pseudowords with high– or low–sublexical frequency components, which we expected would reveal areas related to the generation of articulatory codes but not areas related to phonological encoding. We found significant activation of a premotor network consisting of the dorsal precentral gyrus, the inferior frontal gyrus bilaterally, and the supplementary motor area for low– versus high–sublexical frequency pseudowords. Based on our hypothesis, we concluded that these areas and in particular the LIFG are involved in phonetic and not phonological encoding. We further discuss our findings with respect to the mechanisms of phonetic encoding and provide evidence in support of a functional segregation of the posterior part of Broca's area, the pars opercularis.

Keywords: articulation, fMRI, left inferior frontal gyrus, pars opercularis, phonological processing

Introduction

Even though Broca's area has been associated with speech and articulation since the 19th century, the exact role that it plays in the process is still a matter of debate. Characteristically, in recent models on the neuroanatomy of language, Broca's area has been associated with quite different processes. In one viewpoint, Indefrey and Levelt (2004) hypothesized that Broca's area was engaged at the level of phonological processing and was particularly associated with the process of syllabification. In contrast, in a model proposed by Hickok and Poeppel (2004), Broca's area was assigned to phonetic encoding and implementing the mechanism of retrieving or generating the articulatory codes. In the present study, we try to address this issue and examine whether the left inferior frontal gyrus (LIFG) is involved in the phonological or the phonetic level of language processing. We used event-related functional magnetic resonance imaging (fMRI) and manipulated the phonological properties of pseudowords in a way that separates the processes of phonological and phonetic encoding. This manipulation allowed us to identify the key areas involved in the 2 levels of encoding and to disambiguate the function of Broca's area with respect to these 2 levels.

The processes that lead to the generation of an articulatory-motor plan are a matter of debate amongst researchers (Goldrick and Rapp 2007). However, it is commonly accepted that syllabic, metrical, and featural information is specified in a phonological representation prior to the generation of the motor plan (Levelt 1999). In extended reviews of studies on word production by Indefrey and Levelt (2000, 2004), it was suggested that in the final stages prior to phonetic encoding and the generation of the articulatory representation, the phonological code of a given word is spelled out into its different phonemic segments, incrementally clustered into syllables, and assigned a metrical structure. As syllables are created, they are then rapidly turned into sequences of motor gestures, also known as gestural scores (Browman and Goldstein 1988).

In this account of word production, it is assumed that there is a different mechanism for dealing with high- and low-frequency syllables. Based on the notion that speakers tend to reuse only a small number of syllables and on evidence that pseudowords with high-frequency syllables are faster to produce than their low-frequency counterparts (Cholin et al. 2006), it was proposed that the articulatory scores for frequent syllables are precompiled and stored in a repository called the "mental syllabary" (Levelt and Wheeldon 1994). In contrast, the articulatory representations for less-frequent syllables are compiled online (Levelt et al. 1999).

Neuroanatomically, the processes of generating lexical phonological representations have been associated with 2 regions: the middle and posterior superior temporal gyrus (STG), also known as Wernicke's area (Fiez et al. 1999; Indefrey and Levelt 2000; Hickok and Poeppel 2004), and Broca's area, specifically the pars opercularis, roughly corresponding to Brodmann area (BA) 44 (Poldrack et al. 1999; Burton et al. 2000; Indefrey and Levelt 2000). The latter region in particular has been shown to facilitate sublexical processes that require explicit segmentation, such as tasks where subjects perform phonological decisions like phoneme monitoring, phoneme discrimination, or phoneme sequencing (Zatorre et al. 1992, 1996; Demonet et al. 1996; Poldrack et al. 1999; Burton et al. 2000). In the proposed model by Indefrey and Levelt (2004), the LIFG is part of a network related to syllabification, whereas the premotor cortex (BA6) is responsible for compiling and storing the motor codes for the individual syllables, that is, it is the location of the mental syllabary (Levelt and Wheeldon 2004).

In recent review papers, Hickok and Poeppel (2004, 2007) proposed a different model for understanding linguistic processing and the role of the LIFG. Inspired by the theory of the "mirror neuron system" and the idea of sensory–motor integration (di Pellegrino et al. 1992; Rizzolatti and Arbib 1998; Rizzolatti and Craighero 2004), they hypothesized that there is a common interface between speech perception and production. This interface also facilitates phonemic-to-articulatory code translation and supports a "motor theory of speech perception" (Liberman and Mattingly 1985). Broca's area is part of the sensory–motor integration interface, and in this sense, it is directly involved in the generation or retrieval of the articulatory codes. Following a computational model of speech production, the proposed role of the posterior Broca's area (along with the ventral premotor cortex) is to hold a "speech sound map," that is, representations of phonemes or frequent syllables and their associated motor programs (Guenther et al. 2006).

The concept of the speech sound map is similar to that of the mental syllabary presented by Indefrey and Levelt (2004). Where the 2 theories differ is the role of the posterior part of Broca's area. According to Hickok and Poeppel (2000, 2004, 2007), Broca's area is involved in phonetic encoding and the generation of the articulatory scores because it serves as a store for articulatory representations. On the other hand, according to Indefrey and Levelt, the role of Broca's area is to support syllabification and postlexical phonological processing, that is, processes that are a step before the retrieval or compilation of the articulatory codes.

In this study, we investigated the role of Broca's area in generating an articulatory-motor plan. We specifically wanted to address whether the posterior part of Broca's area (pars opercularis) is involved in phonological processes, such as syllabification, or in directly retrieving or compiling the articulatory gestures. To do this, we used event-related fMRI to monitor the changes in blood oxygenation while subjects performed a delayed pseudoword repetition task. The presented stimuli differed in length (4 vs. 2 syllables) and sublexical frequency of segments and syllables (low vs. high sublexical frequency). We anticipated that we would be able to identify 1) the regions involved in phonetic encoding and 2) disambiguate the role of the pars opercularis in single-word production. Specifically, if Broca's area is involved in syllabification and phonological processing prior to the encoding of the articulatory scores, it would only show a strong effect of length, but not sublexical frequency. On the other hand, if Broca's area is the site of the mental syllabary, we expected to see significant effects of both length and frequency manipulations.

Materials and Methods

Subjects

Fifteen healthy, monolingual native speakers of American English were chosen to participate in the study (8 males and 7 females) with mean age of 26 years (range = 20–35). Two subjects (1 male and 1 female) were excluded from analysis because of excessive head motion. All the volunteers reported that they were right handed, with normal hearing and with no history of previous neurological or psychiatric disease. Volunteers were paid for their participation in the 2-h scanning session, in compliance with the institutional guidelines. Prior to testing, volunteers provided written informed consent as approved by the National Institute on Deafness and Other Communication Disorders–National Institute of Neurological Disorders and Stroke Institutional Review Board (protocol NIH 92-DC-0178).

Table 1
Stimulus features

Condition	Bigram PP	Phoneme PP
4 Syllables, high PP, for example \hɛ.tə.tɛ.sɝ\	0.0251 (±0.0093)	0.4888 (±0.0681)
4 Syllables, low PP, for example \gɔ.fɑ.θo^{w}.jɝ\	0.0013 (±0.0012)	0.1251 (±0.025)
2 Syllables, high PP, for example \kɪ.kɛb\	0.0181 (±0.007)	0.2965 (±0.0427)
2 Syllables, low PP, for example \go^{ɪ}.ʃa^{ɪ}m\	0.0004 (±0.0004)	0.061(±0.0194)

Note: table with examples of the stimuli used in each category (phonetic transcription) and their features. For each category, we include the mean (±SD) PP measures for both biphones and phonemes. Audio samples of the stimuli examples are provided online as Supplementary Material.

Stimulus Materials

Four sets of 36 pseudowords were created (a total of 144 items) varying in length and sublexical frequency: 4-syllable low frequency, 4-syllable high frequency, 2-syllable low frequency, and 2-syllable high frequency. The 4 sets of stimuli consisted of alternating consonant–vowel (CV) biphones plus a final consonant, that is, CVCVC and CVCVCVCVC for 2- and 4-syllable pseudowords, respectively. The 4-syllable pseudowords contained 2 stresses (a primary and a secondary stress). However, the position of the stressed syllables within the pseudowords varied to allow greater flexibility in the creation of the data set and avoiding the creation of ungrammatical syllables. Examples of the stimuli are presented in Table 1 (audio files of the examples are provided online as Supplementary Material). As a measure of length, we chose number of syllables and phonemes, with 2 syllables as the minimum length. Two-syllable pseudowords were preferred over monosyllabic ones to allow better control of phonological neighborhood density, which decreases as the word length increases (Pisoni et al. 1985). As a measure of sublexical frequency, we chose the phonotactic probability (PP) of phonemes and biphones. Phonotactic probability refers to the frequency with which legal phonological segments and sequences of segments (i.e., biphones) occur in a given language (Jusczyk et al. 1994). As observed in the syllable frequency effect, low PP pseudowords have slower response time than high PP ones, reflecting the load in the phonetic encoding process (Vitevitch et al. 1997, 1999; Vitevitch and Luce 1998).

All the syllables, with the exception of 2, that were used to construct the pseudowords were chosen from a corpus of previous linguistic studies on the effects of PP (Vitevitch et al. 1997; Frisch et al. 2000) such that they were rare, but not illegal (in the case of low-frequency items), and that they satisfied our criteria for frequency. The 2 additional syllables that we included were /θo^{w}/ and /θɝ/. Both of these syllables had a biphone probability greater than zero and were included to increase the variability of the generated data set. The PP for each biphone and phoneme was calculated (Vitevitch and Luce 2004), and pseudowords were created such that each pseudoword consisted entirely of high- or low-probability segments (depending on its category).

To reduce the amount of similarity between the stimuli, no 2 syllables occurred in the same pseudoword more than once and no pseudoword appeared as a contiguous part within another pseudoword. All items were further checked for immediate phonological neighbors using a "one phoneme change" rule, that is, no stimulus could be turned into a word by 1) changing one phoneme into another, 2) deleting one phoneme, or 3) adding one phoneme. Even though phonological neighborhood density and PP are correlated, we expected that by controlling for immediate neighbors, the differences in neighborhood density between items with different PP would not be emphasized. Effects related to PP would then be related to phonetic encoding and not phonological word retrieval, which would arise by manipulating phonological neighborhood density (Okada and Hickok 2006). As a result, low- and high-sublexical frequency items differed systematically only with respect to the positional frequency of their phonemes and syllables. Finally, to avoid morphological confounds, any

sequences that ended with a high-probability final rime, for example, /-æs/ and /-æd/, which could be interpreted as inflectional suffixes, were also omitted from the data set.

To record the stimuli, we recruited a female, monolingual American English volunteer. Prior to the recording, the volunteer was trained to pronounce the data set correctly and rehearsed the items a number of times to familiarize herself with the data set. The stimuli were read from a laptop screen and spoken in isolation as naturally and as clearly as possible. All stimuli were recorded in a single session in a nonechoic, sound-attenuated booth. They were digitally recorded using a Shure SM58 vocal microphone at 44.1-kHz sampling rate and were saved at 16-bit resolution. Two or three recordings were made for every stimulus, which were then edited into individual files and screened for accuracy and fluency. The most accurate recording of each item was chosen for the stimulus list. The chosen stimuli were then transcribed, and their segment and biphone PP was recalculated to take into account the cases where there were some differences in the pronunciation. In the resulting lists, the differences between the average segment and the biphone probabilities over both 4- and 2-syllable pseudowords were statistically significant (phonemes: $F_{1,286}$ = 920.2, $P < 0.001$; biphones: $F_{1,286}$ = 763.9, $P < 0.001$). Higher frequency pseudowords had higher PP scores than lower frequency pseudowords (see Table 1 for more details on the category PP).

Experimental Design and Procedure

Thirty-six items per condition were presented over the course of 2 experimental fMRI runs. Each item was presented to the subject auditorily using an fMRI compatible (pneumatic) system for auditory delivery (Avotec SS-3100, Silent Scan system). After a delay of 6 s, a probe (1 of 2 versions of a bell sound) was heard instructing the subject to repeat the presented pseudoword either overtly or covertly (depending on the type of probe). During the delay period, the subjects were given specific instructions to rehearse the presented stimulus covertly. They did not know prior to the presentation of the relevant probe whether they would be asked to respond overtly or covertly, and so we expected that they would fully retrieve the articulatory scores for the presented pseudoword. Each trial lasted 8 s (Fig. 1*A*).

Stimulus presentation was in a pseudorandom, fast event-related fashion, whereby the order of occurrence for the conditions was controlled by a combination of 3 binary shifted versions of an m-sequence (one shifted by 9 bins and the other by 18 bins with respect to the first one; see, e.g., Fig. 1*B*). The use of m-sequences (Buracas and Boynton 2002; Kellman et al. 2003) to control stimulus delivery allowed for a simple and efficient way to increase design efficiency and minimize the chance of significant correlation between the regressors, even in case of post hoc exclusion of incorrect trials. The binary m-sequence used in the study had a length of 63 bins (corresponding to the number of trials per run) and was padded in the beginning with 9 more trials, which were not analyzed for the purposes of this study. The purpose of these onset trials was to allow for the subject to get comfortable with the task and the noisy environment in the scanner.

Prior to the onset of the experiment, all subjects performed a 150-min practice session outside the scanner to allow them to become familiar with the structure of the task and its demands. The material used as the training set (10 items per category) contained pseudowords with features similar to the ones presented during the experimental runs but from an unrelated set (built from different syllables) to avoid habituation and familiarity.

Because of the concern that, during the scanning session, the scanner noise would mask out some of the stimuli, a quality check run was performed prior to the onset of the experimental runs. During this run, a set of pseudowords (not used for the experimental set but recorded in the same session as the experimental set, i.e., with the same amplitude and recording characteristics) was presented to the subject. The volume of the headset was then adjusted based on the subject's feedback to ensure protection from exposure to a noisy environment, comfort, and clear stimulus delivery. Images acquired during this test run were also submitted to a quality check to make sure that they were free from artifacts.

During the scanning session, subject responses were recorded using a dual-channel, noise canceling, fiber optic microphone (Dual-Channel Phone-Or by Optoacoustics Ltd, Or-Yehuda, Israel). This system is specifically designed for use in magnetic resonance imaging (MRI) environments and offers real-time adaptive elimination of the MRI acoustic noise from the signal. This allowed us to record both the subject responses and the timing of their responses. However, due to concerns that the filtering algorithm introduced a small, random delay in the recording of the responses, we did not consider the estimates of the subject response timing reliable. Thus, as a behavioral measurement, we only used subject response accuracy.

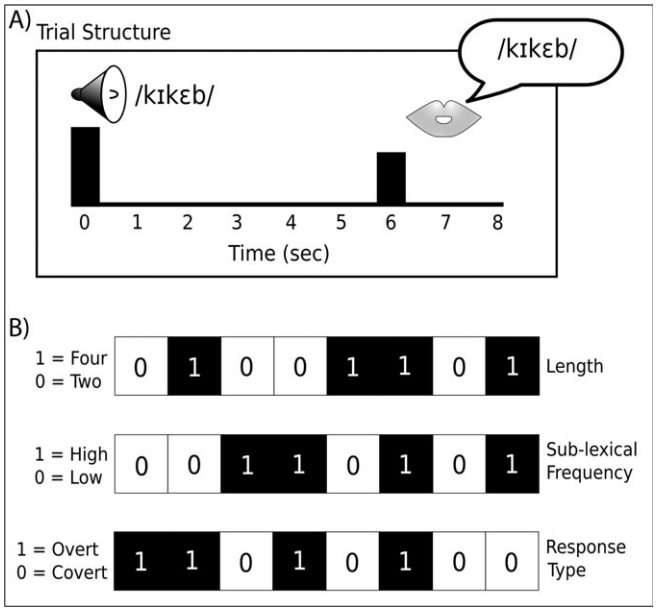

Figure 1. During the experiment, subjects were asked to listen to pseudowords and to repeat them either overtly or covertly after a 6-s delay. The structure of each trial is shown in (*A*). The stimulus is presented auditorily at 0 s and then subjects wait for the response probe. During the delay period, they are instructed to covertly rehearse the stimulus and are not aware of the type of response (overt or covert) before they hear the probe. The type of stimulus that will be presented in each trial is determined pseudorandomly by a combination of 3 m-sequences. In (*B*), we present an example of 3 binary sequences that resemble those used in the experiment. Each sequence is associated with an experimental factor. In the example provided, the top sequence controls the length of the stimulus (1 for 4 syllables and 0 for 2 syllables), the middle sequence controls sublexical frequency (1 for high and 0 for low), and the bottom sequence controls response type (1 for overt and 0 for covert). For example, the combination 0 1 0 would retrieve a 2-syllable, high-frequency pseudoword and the covert response probe.

fMRI Data Acquisition

Imaging was performed on a 3.0-T MRI system (General Electric, Milwaukee, WI), equipped with Cardiac Resonance Module whole-body gradients. For improved signal-to-noise ratio (SNR) and higher spatial resolution, we used a custom-built 16-channel MRI receive array (Nova Medical, Wilmington, MA; de Zwart et al. 2004) connected to a custom-built 16-channel MRI receiver. For the functional scans, we used single-shot, rate-2, sensitivity-encoded (SENSE), gradient-echo, echo-planar imaging (EPI) (de Zwart et al. 2002). A total of 32 axial slices were acquired interleaved (time echo [TE] = 31 ms, flip angle of 90 degrees, time repetition [TR] = 2 s, and acquisition bandwidth 250 kHz) with an in-plane resolution of 2.3 × 2.3 mm^2 (96 × 72 matrix, 22.4 × 16.8 cm^2 field of view [FOV]) and slice thickness = 2 mm (gap = 0.3 mm). Four volumes were acquired during each trial. The combination of the dedicated receive array with SENSE EPI allowed a 2- to 4-fold improvement in SNR and a 50% reduction in geometric distortions relative to a conventional setup with a birdcage head coil (de Zwart et al. 2004). The reduced geometrical distortions of SENSE EPI are due to its use of a shortened data acquisition window compared with conventional EPI at the same spatial resolution.

To increase the efficiency of subject motion correction, we acquired isotropic voxels (2.3 mm cube side). However, the resulting smaller-than-usual thickness of the slices put a constraint on the brain volume that could be imaged. We did not have a hypothesis about the involvement of any areas below the superior temporal sulcus (STS), and we therefore acquired images in a slightly oblique position, covering an area from below the STS to the top of the head. By avoiding the lower parts of the cortex (e.g., the inferior temporal areas), we also avoided geometrical distortions and artifacts that are caused by articulatory muscle movement (Birn et al. 2004). To facilitate slice selection, a sagittal 2-dimensional anatomical image was acquired prior to the onset of the functional runs. This image was inspected for specific anatomical landmarks such as the anterior commissure and the STS and was used to make the slice selection. At the end of the scanning session, high-resolution spin-echo T_1 anatomical images were acquired at the same location as the functional EPI scans. The scanning parameters for the anatomical image were as follows: TR = 700 ms, TE = 13 ms, 256 × 192 data matrix with a 22.4 × 16.8 cm^2 FOV, resulting in 0.86 × 0.86 mm^2 in-plane resolution, and 2 mm slice thickness (with 0.3 mm gap).

To minimize head movement during the scanning sessions, we used head padding and a velcro strap, mounted on each side of the head coil and positioned on the subject's forehead at the line just above the eyebrows. The purpose of the strap was to act as a motion reference point for the subject. Head movement, especially in the z (head–foot) direction, would cause a strain on the strap, make the subject aware of the movement and cause him/her to restrict it and return to the original position. Prior to the onset of the scanning session, the subjects were given instructions about how to restrict their head movement and about the function of the velcro strap. Tests were also performed to ensure that the strap was properly placed, and the subjects could feel it when moving during speech.

Image Preprocessing

All analyses and image preprocessing were carried out using the SPM5 software package and associated toolboxes (http://www.fil.ion.ucl.ac.uk/spm/software/spm5). Preprocessing included slice-timing correction and an optimized motion correction routine to ensure good quality registration (Oakes et al. 2005). Images were then registered to the Montreal Neurological Institute (MNI) anatomical template and transformed into MNI stereotactic space to allow for group comparisons. The functional data were then smoothed with an isotropic Gaussian filter kernel of 6 mm (full width at half maximum) to improve SNR.

To quantify the effect of subject movement on the quality of our data, we inspected the data using the ArtRepair toolbox for SPM5 (Mazaika et al. 2007) and examined the realignment parameters provided by the SPM5 motion correction procedure. We were particularly interested in scan-to-scan (incremental) motion during the task, that is, the change in position between the image acquired during the subject response and its immediate preceding image. In previous studies on speech-related motion (Barch et al. 1999), it was shown that speech-related motion is mainly scan-to-scan motion affecting the first scan acquired after the response probe. To assess the effects of speech-related motion on our data, we performed a 3-factor analysis of variance (ANOVA) with within-subject factors response type, stimulus length, and sublexical frequency and dependent variable the 6 motion estimates for incremental (scan-to-scan) movement. The analysis revealed a significant main effect of response type in all directions ($F_{1,12} > 26$, $P < 0.004$ for all directions). In agreement with other studies (Barch et al. 1999; Shuster and Lemieux 2005), the incremental movement was overall quite small and greater for overt response trials (mean ± standard deviation [SD] displacement was 0.039 ± 0.014 mm for translations and 0.034 ± 0.012° for rotations) than covert response ones (mean ± SD was 0.02 ± 0.008 mm for translations and 0.017 ± 0.006° for rotations).

Additional significant effects were present for length in the pitch rotation and for both the main effect ($F_{1,12} = 5.9$, $P < 0.04$) and the interaction between length and response type ($F_{1,12} = 19$, $P < 0.001$). Four-syllable pseudowords (mean ± SD pitch displacement was 0.038 ± 0.020°) produced greater movement than 2-syllable pseudowords (mean was 0.034 ± 0.016°) especially during overt responses. Finally, in the y direction, there was a significant main effect of sublexical frequency ($F_{1,12} = 6.3$, $P < 0.03$) and interaction between sublexical frequency and response type ($F_{1,12} = 10.8$, $P < 0.01$). Low-frequency items caused greater movement (mean ± SD 0.021 ± 0.013 mm) than high-frequency items (0.019 ± 0.010 mm), especially during overt response trials. To remove effects related to subject movement, we included the realignment parameters in the design matrix as effects of no interest. Finally, we also inspected the movement parameters for extreme movement. We took into account both incremental movement and absolute movement (the displacement of a scan with respect to the realignment reference scan of the time series, i.e., in our case, the first image in the time series). Our criteria for inclusion in the study were that a subject would not show absolute motion greater than the voxel size and incremental motion greater than 1 mm in translations and 1° in rotations. All subjects met the absolute motion inclusion criteria, but not the incremental motion. Two subjects showed movement greater than our criteria and were consequently excluded from the analysis.

Further examination using the ArtRepair toolbox revealed that in a few cases, incremental movement even as low as 0.5 mm induced global signal changes greater than 1.5% of the mean and "stripe-like" artifacts on the image. To ensure the quality of our data and to completely remove their effect from the analysis, we also added an additional regressor for images that showed changes in the global signal greater than 1.5% of the mean followed by a greater than 0.5 mm incremental movement (Mazaika et al. 2007).

Behavioral Data Analysis

In order to get an estimate of subject performance and ensure that the subjects were performing the task as instructed, we estimated the subject response accuracy. To calculate it, we monitored and phonologically transcribed all subject responses. However, because of the low quality of the recording, resulting from the noise reduction filtering, a precise phonetic transcription of the subject response was not always possible and the nearest phonological transcription was used. Cases where the recording was unintelligible because of noise were not included in the analysis. The resulting transcriptions were compared with the target stimulus phoneme-by-phoneme, and a score was calculated based on the number of correct phonemes (token count). If a phoneme was omitted in the subject response, it was scored as a mismatch, for example, if the target was /kɪkeb/ and the response was /keb/, the first 2 phonemes were counted as a mismatch and the final phonemes were counted as a match. To determine a match between the target and the response, we used broad phonemic criteria and ignored differences between allophones (Vitevitch and Luce 2005). The scores were then submitted to a 2-way ANOVA with factors length and sublexical frequency.

Even though we were not able to extract a very detailed phonetic transcription, our interpretation of the data does not dependent on the subtle phonetic details of the subjects' performance, for example, distinguishing between 2 allophones. The primary reasons for analyzing the behavioral results were to identify incorrect trials, to ensure that the subjects were performing the task as instructed, and that the difference between low- and high-sublexical frequency items was retained in the subject response. For this purpose, we also estimated the PP of the subjects' overt responses in the same way as we did for the stimuli (Vitevitch and Luce 2004). To determine whether there is a significant difference between the 2 conditions, we performed a paired t-test. Finally, we also examined the subject recordings to identify trials that were incorrectly answered (i.e., responses on covert trials or no response on overt trials). These trials were included to a regressor of no interest and excluded from the fMRI data analysis.

fMRI Data Analysis

Statistical analysis of the factorial event-related experiment was performed using SPM5. The hemodynamic response function (HRF) for each trial was modeled using a finite impulse response function (FIR) with 12 bins (duration of 2 s) to capture the temporal components of a delayed response task. Stimulus presentation was modeled as a delta function. A 2-way, random-effects, within-subject

ANOVA with factors length (4- vs. 2-syllable pseudowords) and sublexical frequency (low vs. high) was performed. Each of the 4 different resulting types of trials, for example, 4-syllable and low sublexical frequency, was modeled by separate regressors, and the main effects and interactions were evaluated by contrasting within or across (interactions) the levels of each factor. To perform group statistics, the contrast images for each effect and for all subjects were submitted to a 1-way ANOVA (with 12 levels). *T*-contrasts testing for the predicted shape of the HRF (a canonical, 2 gamma function; Friston et al. 1998) were performed to produce maximum intensity projections and reveal voxels whose differential activity pattern conforms to the shape of the HRF. SPMs were thresholded at $P < 0.001$ uncorrected at the voxel level and $P < 0.05$ corrected for familywise error (FWE) at the cluster level (Hayasaka and Nichols 2003). For our study, significant clusters had on average more than 85 voxels.

In order to analyze the contrast estimates for the LIFG, we used the cytoarchitectonic probability map for left hemisphere BA44 (Eickhoff et al. 2005). For each of the main effects of interest (length, frequency, and response type), we identified the voxels within the activated clusters that were part of BA44. We then extracted the average beta weights (over cluster voxels) for each of the 4 conditions of interest in the design (4-syllable low frequency, 4-syllable high frequency, 2-syllable low frequency, and 2-syllable high frequency) and for all subjects. A single value corresponding to the weighted sum of the estimates across the FIR (weighted by the HRF) was then extracted for each of the 4 conditions and subjects and used in multiple 2-sided *t*-tests testing for effects of frequency, length, or the difference between the 2 conditions within each region. This approach followed the implementation of random-effects analyses in the Marsbar SPM toolbox (Brett et al. 2002). Significance was determined using a threshold of $P < 0.05$. Where appropriate (more than 1 region of interest [ROI]), the *P* values were adjusted to correct for multiple comparisons (Bonferroni correction).

To ensure that the significant activations observed during the delay period for both the whole-brain and the LIFG analyses were not related to subject motion, we extracted and inspected the parameter estimates for each significantly activated cluster over the window of the FIR (24 s). The time course of movement-related activations is very different from that of blood oxygen level–dependent (BOLD) related activations. Whereas motion-related signal changes appear as large spikes in the signal intensity for the first few images at the time of the subject movement, BOLD-related signal changes follow a curve similar to the HRF (Birn et al. 1999). It should also be noted that significant effects for length and frequency were estimated over both covert and overt responses, and so we expected that the contribution of motion-related artifacts to the significant activations observed would be minimal, if any.

Results

Behavioral Results

To test for effects of length or frequency on subject performance, we measured subject response accuracy. Based on previous results, we expected to find a decrease in response accuracy for low-frequency pseudowords, but we did not expect to find an effect of length. We performed a 2-way ANOVA with within-subject factors: length and sublexical frequency. As expected, we found that there was a significant main effect only for frequency ($F_{1,12} = 14.6$, $P < 0.003$). No other main effects or interactions were significant. Mean (±SD) accuracy rates were 64.5% (±15) for low-frequency pseudowords and 75% (±13) for high. The relatively low accuracy scores were expected, considering the nature of the task (pseudoword repetition) and the noisy environment. All subjects' performance accuracy was within 3 SDs of the group mean (70%, SD = 13).

Finally, to verify that there is a significant difference in sublexical frequency between the responses, we calculated the phoneme and biphone PP of the subjects' overt responses and performed a 2-sided *t*-test to compare high- versus low-frequency responses. For both biphone and phoneme measurements, the differences were significant ($t_{12} = 14.66$, $P < 0.001$, for biphones and $t_{12} = 15.74$, $P < 0.001$, for phonemes). Mean (±standard error [SE]) PP scores for high-frequency responses were 0.0193 (±0.0009) for biphones and 0.3656 (±0.0145) for phonemes. Low-frequency PP scores were 0.0025 (±0.0006) for biphones and 0.1187 (±0.0091) for phonemes. From the above results, we can conclude that the subjects perceived the differences between low- and high-frequency targets and performed the task according to the instructions.

fMRI Results

Phonological Encoding

To map the areas involved in phonological encoding, we compared the activation levels invoked for processing 4- versus 2-syllable pseudowords (over both low- and high-frequency syllables). A significant main effect of length (4- greater than 2-syllable stimuli) was observed in a large perisylvian network extending bilaterally across the STG, the precentral gyrus (PrCG), and the supplementary motor area (SMA), as well as the LIFG (cf., Fig. 2*A* for whole-brain results and Fig. 2*C* for significantly activated voxels within the LIFG). The largest activations were observed in the left hemisphere for a cluster that covered both the PrCG and STG. In particular for the STG, the cluster covered a large portion of the middle and posterior STG including the upper banks of the STS and an area in the junction between the parietal and the temporal lobe also referred to as the Sylvian parietotemporal area (SPT; cf., Table 2 for the coordinates of the significantly activated areas). The left STG (LSTG) has been previously implicated in phonological processing (Indefrey and Levelt 2000, 2004; Graves et al. 2007), whereas the left PrCG is a known premotor area and as such it has been associated with phonetic encoding. A similar effect could also be observed for the LIFG. The activated area was located on pars opercularis and ran along the inferior frontal sulcus (IFS). In accordance to our hypothesis, we expected that both phonological and phonetic encoding processes would show an effect of length. What distinguishes the 2 processes is their sensitivity to sublexical frequency. If a region is involved in phonological processing, we would not expect it to show significant sublexical frequency effects. On the other hand, if it is, we would expect it to show significant effects for both conditions, length and sublexical frequency.

Phonetic Encoding

Comparing pseudowords with low versus high PP syllables and segments revealed regions that showed an effect for sublexical frequency. Based on our hypothesis, areas that showed a frequency effect reflect the process of phonetic encoding, that is, articulatory code generation (Indefrey and Levelt 2000). Four regions showed significant main effects of frequency: the left hemisphere dorsal PrCG, the left hemisphere SMA (LSMA), and the inferior frontal gyrus (IFG) bilaterally (cf., Table 2 for a detailed list of the activated regions and Fig. 2*B* for a map of the significantly activated areas). Activity in the LSTG did not reach significance ($P < 0.2$ cluster size, FWE corrected).

We also tested for the opposite contrast, high- versus low-frequency pseudowords in order to see whether the areas

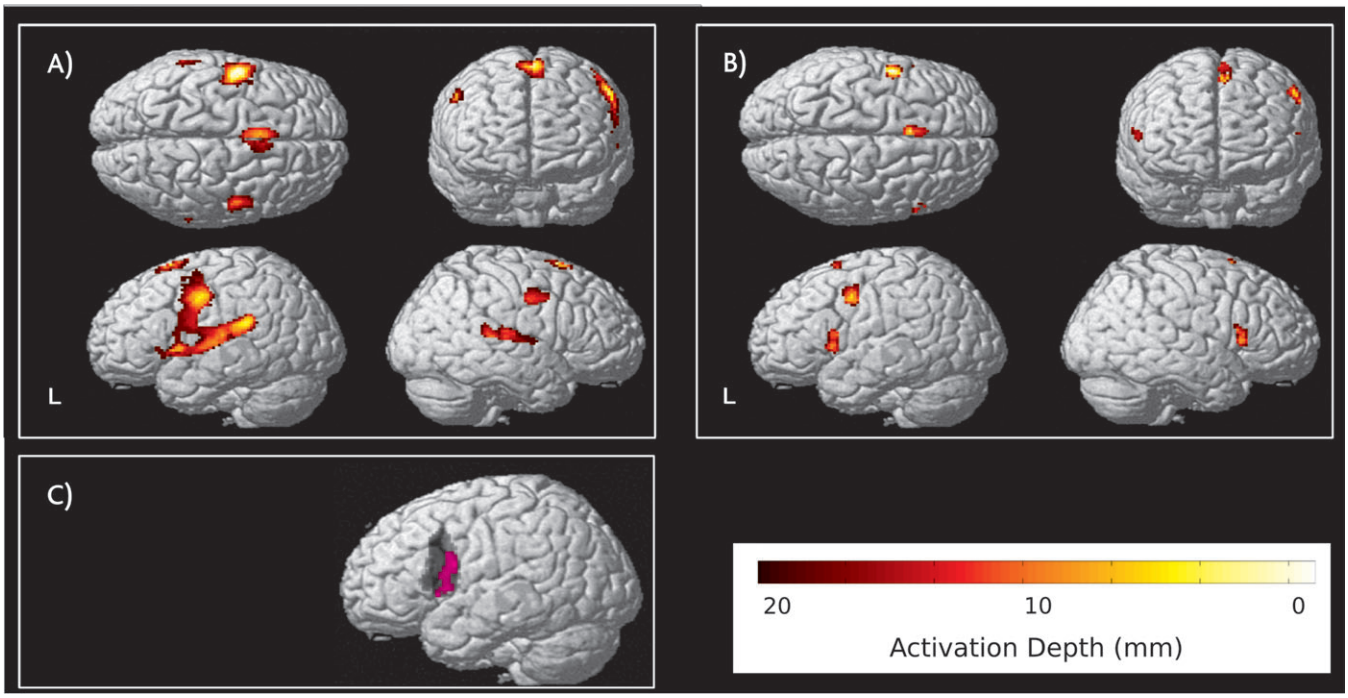

Figure 2. Surface renderings of significant activations in the whole-brain group analysis for length (*A*) and sublexical frequency (*B*). In (*A*), an extended perisylvian and premotor activation including the LIFG showed significantly higher activation for 4 versus 2 syllables. In (*B*), premotor areas including the dorsal PrCG and the IFG bilaterally showed significantly higher activation for low- versus high-frequency pseudowords. In (*C*), we show the main effect of length within left BA44 (significantly activated voxels appear in magenta) using a small volume correction approach (SVC). BA44 (shaded area) was defined using a cytoarchitectonic probability map of the area (Eickhoff et al. 2005). Maps are thresholded voxelwise at $P < 0.001$ uncorrected and clusterwise at $P < 0.05$ FWE corrected. Color grading in (*A*) and (*B*) reflects depth, with brighter voxels on the surface. The maximum depth of the projected voxels is 20 mm. L, sagittal view of the left hemisphere.

associated with retrieving high-frequency, precompiled syllables from the mental syllabary are different from the ones associated with online generation of articulatory scores. No areas showed higher activation for high- versus low-frequency syllables. There were also no significant interaction effects between length and sublexical frequency.

Left IFG

To further test our hypothesis about the involvement of Broca's area in phonetic processing, we performed an ROI analysis. A region corresponding to left hemisphere BA44 (center of mass $x = -53$, $y = 12$, $z = 19$, size = 1160 voxels) was defined using a cytoarchitectonic probability map of area BA44 (Eickhoff et al. 2005). In a random-effects 2-way ANOVA with factors length (4 vs. 2 syllables) and sublexical frequency (low vs. high), the LIFG showed a main effect for both factors ($t_{12} = 1.97$, $P < 0.04$, and $t_{12} = 2.56$, $P < 0.02$, for length and frequency, respectively).

Because the LIFG showed effects for both length and frequency, we further investigated whether there were any signs of functional segregation within the IFG and in particular the pars opercularis, as had been observed in other studies (Molnar-Szakacs et al. 2005). For the 2 conditions, length and frequency, we observed 2 clusters within the LIFG, which were only partly overlapping (9 voxels out of 82 and 79, respectively, for the 2 clusters; Fig. 3). The distance between their center of mass was 9 mm, that is, a factor of 1.5 greater than the smoothing kernel (6 mm), with the cluster showing a greater effect of length following the anterior banks of the IFS and extending more lateral, posterior, and dorsal to the cluster showing a greater effect of frequency. We will refer to the cluster identified during the length condition as dorsal pars opercularis (dPOp) and the cluster identified for the frequency condition as ventral pars opercularis (vPOp) because of their anatomical differences and in agreement with previous evidence.

Both the dPOp and the vPOp exhibited effects of frequency and length, though the frequency effect for dPOp was just

Table 2
Brain regions modulated by length and frequency

Contrast	Region	Coordinates			*T*	No. of voxels
		x	*y*	*z*		
4 > 2 Syllables	Left PrCG	−56	−4	44	7.87	2097
	LSTG[a]	−60	−12	4	6.76	
	Left SPT junction[a]	−56	−38	20	5.82	
	LIFG[a]	−60	4	20	4.63	
	LSMA	−4	10	68	7.21	388
	Right STG	50	−22	8	5.45	393
	Right SPT junction[a]	64	−32	10	5.24	
	Right PrCG	50	−4	40	5.30	176
Low > high frequency	Left PrCG	−52	2	40	4.77	138
	LSMA	−4	14	58	4.51	122
	LIFG	−54	12	12	4.01	119
	Right IFG	50	18	4	4.23	97

Note: regions significantly activated in the group analysis ($t_{144} > 3.1$, $P < 0.05$ FWE corrected for cluster size). Displayed are the contrasts, the coordinates for the voxels of greatest activity within the activated clusters in MNI stereotaxic space, an anatomical description of the region, the *T* value, and the number of significantly activated voxels.

[a]In the case of very large clusters, multiple peak voxels are reported. They are clustered together with the last entry to include number of voxels.

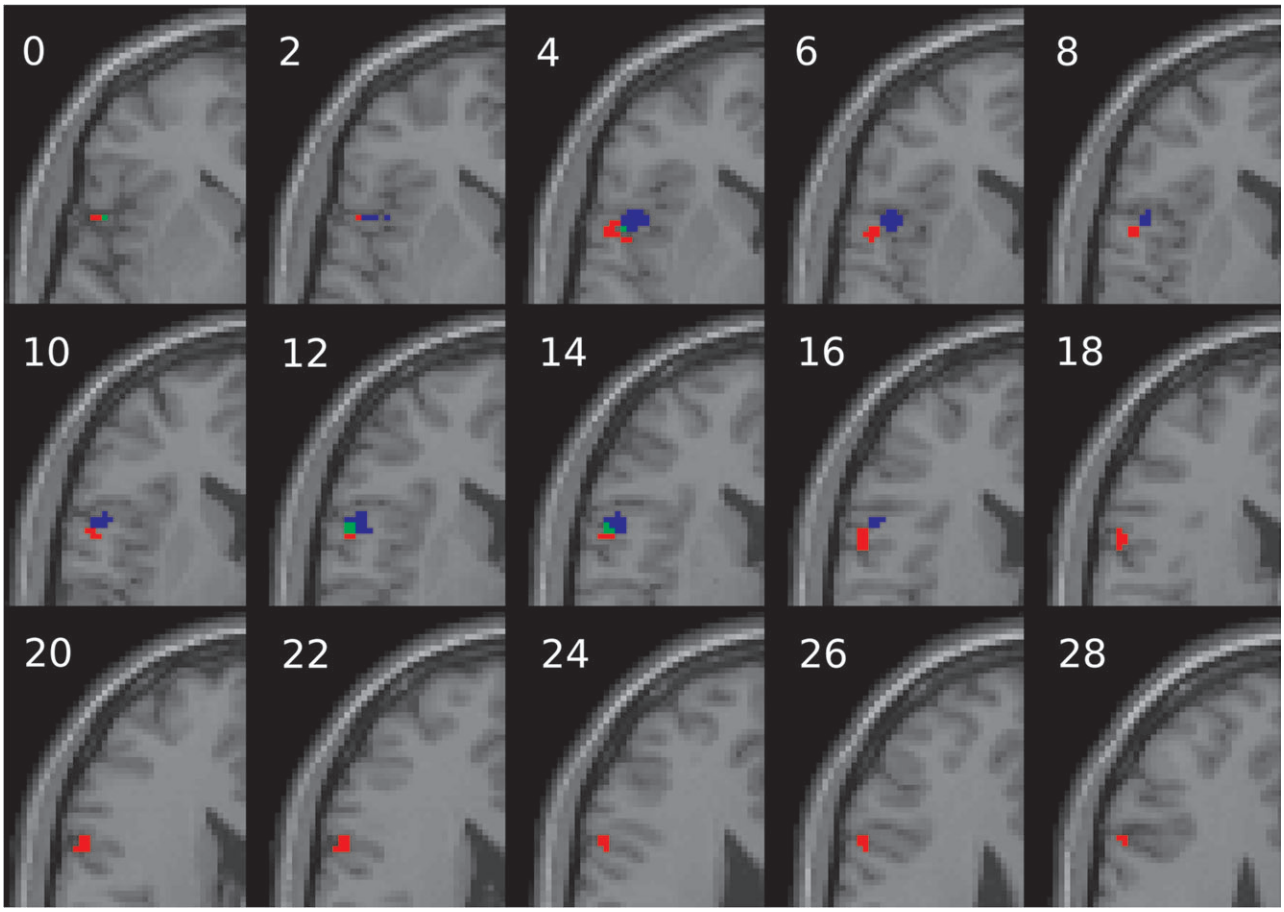

Figure 3. Significant activations within left hemisphere BA44 as defined by a cytoarchitectonic probability map of the area (Eickhoff et al. 2005). Shown in red are voxels significantly more activated for 4 versus 2 syllables. This cluster extends from $z = -2$ (slice not shown) to $z = 28$. The largest effect for length is located dorsally, at [−60 4 20]. Shown in blue are voxels significantly more activated for low versus high sublexical frequency. The largest effect for frequency is located at [−54 12 12]. Finally, shown in green are voxels that are overlapping for both conditions (size of overlap = 9 voxels). Activations are thresholded at $P < 0.001$ uncorrected voxelwise and $P < 0.05$ FWE corrected clusterwise. *Z* coordinates are in MNI space.

slightly above significance (dPOp frequency: $t_{12} = 2.5$, $P < 0.06$; vPOp length: $t_{12} = 3.2$, $P < 0.02$ corrected for 2 ROIs). This difference already suggests that there might be a functional segregation within the pars opercularis of the LIFG. To further examine whether there is a functional difference in the activation between the 2 clusters, we examined the region (dPOp vs. vPOp) by experimental condition (length vs. frequency) interaction (Friederici et al. 2006). We performed a 2-sided paired *t*-test on the region-specific differences between the length and the frequency conditions and found a significant region-by-condition interaction ($t_{12} = 3.1$, $P < 0.01$), indicating that there is a robust difference between the 2 clusters in terms of their response to length and sublexical frequency effects. DPOp shows greater activation for length rather than sublexical frequency (mean ± SE length over frequency difference is 0.093 ± 0.051), whereas in vPOp, there is almost no difference between the levels of activation for the 2 conditions (mean ± SE length over frequency difference is 0.002 ± 0.026).

Discussion

In this study, we were able to delineate the cortical areas involved in the phonemic-to-articulatory translation that is necessary for the generation of articulatory codes. By directly contrasting targets with varying length, we manipulated the load on the system of postlexical articulatory-motor production and were able to identify a number of key regions underlying articulation and the overall process of transforming phonological word forms to articulatory codes. In summary, these regions included bilateral (although strongly left lateralized) mid and posterior superior temporal and frontal regions, the premotor cortex, and the SMA. These results are in agreement with current models on word production that describe a left-lateralized, perisylvian network (Indefrey and Levelt 2000, 2004; Hickok and Poeppel 2004, 2007).

To further identify the roles of the different components of the network and in particular to resolve the conflict on the role of the LIFG, we probed the network by manipulating sublexical frequency. Our hypothesis was that only regions that are directly involved in phonemic-to-articulatory translation would show an effect for frequency manipulation. Targets with components of different sublexical frequency (high vs. low) are processed differently (Guenther et al. 2006). High-frequency clusters are precompiled and their articulatory codes are retrieved, as suggested by the fact that they are processed faster than the ones with less-frequent components (Vitevitch and Luce 1998, 2005). The latter are thought to be

compiled online on a segment-to-segment basis (Guenther et al. 2006).

In our experiment, we identified 4 regions that showed an effect related to sublexical frequency (higher activation for low vs. high frequency): the LSMA, the left hemisphere PrCG, and the IFG bilaterally. From previous studies on motor planning and production, it is known that the SMA has a role in motor planning and the preparation of movements. Even though its function is not specifically associated with linguistic processes, it is also part of linguistic motor planning (Riecker et al. 2005). In a recent fMRI study, the pre-SMA was shown to be sensitive to sequence complexity effects both within and beyond the syllable boundaries (Bohland and Guenther 2006). The present findings are in agreement with the current theories on the function of the SMA. The observed frequency effect could simply represent the increased load that is associated with producing new and unfamiliar motor plans (low–sublexical frequency pseudowords) compared with familiar, more rehearsed, and precompiled ones (high–sublexical frequency pseudowords).

The significant activation difference for low– versus high–sublexical frequency pseudowords in the left PrCG is also in agreement with current models on word production (Hickok and Poeppel 2004; Indefrey and Levelt 2004; Guenther et al. 2006). It is worth highlighting that only a small area in the dorsal PrCG was significantly active and that this area has been previously involved in studies examining sensory–motor mapping (Hickok and Poeppel 2004). Hickok and Poeppel propose the existence of a "dorsal stream" in speech processing, which is involved in mapping sound onto articulatory-based representations. The regions that are part of this stream include a posterior inferior frontal area (including Broca's area), a dorsal premotor site, and area SPT (Hickok et al. 2003). The latter region, which lies within the boundaries of the planum temporale, is traditionally associated with acoustic and phonological processing, as well as speech production as the interface for the sound-to-gesture transformation.

In our study, we found that the STG bilaterally shows a greater effect for target length, though the results are strongly left lateralized, and in the left hemisphere, particularly, the effect extends further in the posterior direction to area SPT (Fig. 2*A*). Bilateral STG activation has been observed during both speech perception and production and reflects the processing of the acoustic and phonological properties of the target stimulus (Hickok and Poeppel 2004). This is in contrast to area SPT, which is thought to be involved in translating between acoustic and motor representations. However, in the current study, both STG and area SPT show a similar behavior and a significant main effect for length only and not for sublexical frequency. Therefore, these findings raise doubts on the role of SPT as an auditory–motor interface and suggest that its role is not that different from the rest of the STG, that is, it could also be involved in phonological processes, such as syllabification and segmentation. This claim would be in agreement with initial claims made by Indefrey and Levelt (2000), whereby a portion of the superior temporal lobe was considered as a possible candidate region for syllabification. Another candidate was the LIFG.

In the current study, we found significant bilateral activation in the IFG. The presence of a sublexical frequency effect in the right IFG was surprising because this region has not been included in any of the neuroanatomical models of speech production previously discussed (Hickok and Poeppel 2000, 2004, 2007; Indefrey and Levelt 2000, 2004). Activation in this region has been previously found during pitch processing and specifically for the integration of accent patterns (Geiser et al. 2008). In the current study, the stress pattern between the 2 categories was controlled, and there were no systematic differences. However, it is possible that the increased processing demands for low–sublexical frequency pseudowords also affected the processing of metrical structure. Further research would be needed to identify the exact nature of the differences.

With respect to the LIFG, the pars opercularis showed consistent effects for both length and sublexical frequency (4 vs. 2 syllables and low vs. high frequency, respectively), as well as evidence of functional segregation. The more dorsal part of the area (dPOp) was modulated by differences in stimulus length, whereas the ventral part (vPOp) was modulated by differences in both length and sublexical frequency. The idea that Broca's area is functionally segregated into its 3 anatomical parts (pars opercularis, triangularis, and orbitalis) is well known and well founded (Bokde et al. 2001; Chein et al. 2002; Devlin et al. 2003; Heim et al. 2007). Recently, however, there have also been claims concerning a functional segregation within pars opercularis (Molnar-Szakacs et al. 2005). In a meta-analysis of imaging studies on imitation and action observation, Molnar-Szakacs et al. (2005) identified 2 distinct foci within the pars opercularis, a dorsal and a ventral one, that serve different functions. DPOp shows mirror neuron properties and is significantly active during both action observation and imitation, whereas vPOp shows only motor properties and is only active during imitation.

In agreement with this segregation, we also identified 2 distinct clusters within the pars opercularis with one extending more dorsally than the other. The more dorsal cluster is located closer to the IFS and the premotor cortex and shows greater activation for length manipulation. The vPOp, on the other hand, shows both a main effect of length and sublexical frequency. In the current study, the dPOp is part of a wider area of activation in the left hemisphere PrCG. Therefore, based on its relation to premotor areas, as well as the fact that it is only active for the length condition, we can conclude that the dPOp is involved in phonological encoding and syllabification as proposed by Indefrey and Levelt (2000, 2004). This role is in agreement with other proposed roles such as sequencing discrete units (Gelfand and Bookheimer 2003) or sublexical processing requiring explicit segmentation (Zatorre et al. 1996; Burton et al. 2000; Chein et al. 2002).

The vPOp on the other hand shows a significant effect of both length and frequency, which is in agreement with a role as the cite of the speech sound map or mental syllabary that has been proposed by Guenther et al. (2006). These results are also partially in agreement with the claims made by Molnar-Szakacs and colleagues, who propose that it holds a form of representation of the motor plans that is communicated to the posterior part of the STS (Molnar-Szakacs et al. 2005). In this account, the vPOp is not the location of the speech sound map but only holds a copy of the articulatory codes. The codes themselves are generated elsewhere. The only other possible candidate in our case would be the dorsal premotor cortex, which also showed a significant effect of sublexical frequency. Based on our results, we cannot exclude either possibility.

Research into the functional segregation of the pars opercularis is still in a preliminary phase. The anatomy of the LIFG is highly variable across subjects (Amunts et al. 1999), which makes it difficult to draw any precise conclusions about the exact anatomical borders of the hypothesized segregation of the pars opercularis based on group-averaged results. For the purposes of this study, we have also described the functional segregation of the region using gross anatomical terms such as ventral and dorsal and only in terms of the group tendency. Future research using higher spatial resolution at the single-subject level will be needed to further verify and specify the exact anatomical features of this functional segregation.

Finally, we also note that we did not find any regions showing significant effects for the inverse contrast high– versus low–sublexical frequency. Based on our hypothesis, we would expect that a significant activation for this contrast would reveal the location of the mental syllabary versus the network underlying articulatory code generation. However, based on the computational model proposed by Guenther et al. (2006), the speech sound map (the equivalent of the mental syllabary) does not just contain precompiled frequent syllables but also motor representations for phonemes, common words, phrases, etc. The speech sound map is therefore involved in both processes, though the online compilation of articulatory codes would be computationally more demanding than the retrieval of precompiled gestural scores. This would explain why we do not see increased activity for high- versus low-frequency stimuli because it would be the same network that is underlying both processes.

To conclude, in this fMRI study, we investigated the process of phonological-to-articulatory translation and the role of the LIFG. Based on our findings, we conclude that the LIFG, BA44 in particular, is functionally segregated into 2 subregions following a dorsal–ventral gradient. The dorsal part seems to be involved at the level of phonological encoding as suggested by Indefrey and Levelt (2000, 2004), whereas the ventral part seems to be involved at the level of phonetic encoding and possibly in the translation between phonemic and articulatory representations as proposed by Hickok and Poeppel (2000, 2004, 2007). This finding is in agreement with recent observations on the functional segregation of the pars opercularis and further clarifies the role of the LIFG in language production.

Supplementary Material

Supplementary material can be found at http://www.cercor.oxfordjournals.org/.

Funding

Neuroinformatics Doctoral Training Centre studentship; UK Engineering and Physical Sciences Research Council; Greek Bakalas Bros Foundation to MP; Intramural Research Program of the National Institute on Deafness and Other Communication Disorders of the US National Institutes of Health; Intramural Research Program of the National Institute of Neurological Disorders and Stroke of the US National Institutes of Health to JAdZ and JMJ.

Notes

We would also like to thank Drs Jason Smith, Jieun Kim, Fatima Husain, David McGonigle, Allen Braun, and Jeff Duyn for their support and helpful comments during the design and execution of the study. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (http://www.ecdf.ed.ac.uk). The ECDF is partially supported by the e-Science Data, Information and Knowledge Transformation (eDIKT) initiative. *Conflict of Interest*: None declared.

Address correspondence to Marina Papoutsi, Centre for Speech, Language, and the Brain, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK. Email: marina@csl.psychol.cam.ac.uk.

References

Amunts K, Schleicher A, Burgel U, Mohlberg H, Uylings HB, Zilles K. 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J Comp Neurol. 412(2):319–341.

Barch DM, Sabb FW, Carter CS, Braver TS, Noll DC, Cohen JD. 1999. Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. Neuroimage. 10(6):642–657.

Birn RM, Bandettini PA, Cox RW, Shaker R. 1999. Event-related fMRI of tasks involving brief motion. Hum Brain Mapp. 7(2):106–114.

Birn RM, Cox RW, Bandettini PA. 2004. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. Neuroimage. 23(3):1046–1058.

Bohland JW, Guenther FH. 2006. An fMRI investigation of syllable sequence production. Neuroimage. 32(2):821–841.

Bokde AL, Tagamets MA, Friedman RB, Horwitz B. 2001. Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. Neuron. 30(2):609–617.

Brett M, Anton J, Valabregue R, Poline J. 2002. Region of interest analysis using an SPM toolbox. Proceedings of the 8th International Conference on Functional Mapping of the Human Brain in Sendai, Japan. Vol. 16. Available on CD-ROM in Neuroimage.

Browman CP, Goldstein L. 1988. Some notes on syllable structure in articulatory phonology. Phonetica. 45(2–4):140–155.

Buracas GT, Boynton GM. 2002. Efficient design of event-related fMRI experiments using M-sequences. Neuroimage. 16:801–813.

Burton MW, Small SL, Blumstein SE. 2000. The role of segmentation in phonological processing: an fMRI investigation. J Cogn Neurosci. 12(4):679–690.

Chein JM, Fissell K, Jacobs S, Fiez JA. 2002. Functional heterogeneity within Broca's area during verbal working memory. Physiol Behav. 77(4–5):635–639.

Cholin J, Levelt WJ, Schiller NO. 2006. Effects of syllable frequency in speech production. Cognition. 99:205–235.

de Zwart JA, Ledden PJ, van Gelderen P, Bodurka J, Chu R, Duyn JH. 2004. Signal-to-noise ratio and parallel imaging performance of a 16-channel receive-only brain coil array at 3.0 Tesla. Magn Reson Med. 51(1):22–26.

de Zwart JA, van Gelderen P, Kellman P, Duyn JH. 2002. Reduction of gradient acoustic noise in MRI using SENSE-EPI. Neuroimage. 16(4):1151–1155.

Demonet JF, Fiez JA, Paulesu E, Petersen SE, Zatorre RJ. 1996. PET studies of phonological processing: a critical reply to Poeppel. Brain Lang. 55(3):352–379.

Devlin JT, Matthews PM, Rushworth MFS. 2003. Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. J Cogn Neurosci. 15(1):71–84.

di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. 1992. Understanding motor events: a neurophysiological study. Exp Brain Res. 91(1):176–180.

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K. 2005. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage. 25(4):1325–1335.

Fiez JA, Balota DA, Raichle ME, Petersen SE. 1999. Effects of lexicality, frequency, and spelling-to-sound consistency on the functional anatomy of reading. Neuron. 24(1):205–218.

Friederici AD, Fiebach CJ, Schlesewsky M, Bornkessel ID, von Cramon DY. 2006. Processing linguistic complexity and grammaticality in the left frontal cortex. Cereb Cortex. 16(12):1709–1717.
Frisch SA, Large NR, Pisoni DB. 2000. Perception of wordlikeness: effects of segment probability and length on the processing of nonwords. J Mem Lang. 42:481–496.
Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. 1998. Event-related fMRI: characterizing differential responses. Neuroimage. 7(1):30–40.
Geiser E, Zaehle T, Jancke L, Meyer M. 2008. The neural correlate of speech rhythm as evidenced by metrical speech processing. J Cogn Neurosci. 20(3):541–552.
Gelfand JR, Bookheimer SY. 2003. Dissociating neural mechanisms of temporal sequencing and processing phonemes. Neuron. 38(5): 831–842.
Goldrick M, Rapp B. 2007. Lexical and post-lexical phonological representations in spoken production. Cognition. 102(2): 219–260.
Graves WW, Grabowski TJ, Mehta S, Gordon JK. 2007. A neural signature of phonological access: distinguishing the effects of word frequency from familiarity and length in overt picture naming. J Cogn Neurosci. 19(4):617–631.
Guenther FH, Ghosh SS, Tourville JA. 2006. Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang. 96(3):280–301.
Hayasaka S, Nichols TE. 2003. Validating cluster size inference: random field and permutation methods. Neuroimage. 20(4): 2343–2356.
Heim S, Eickhoff SB, Ischebeck AK, Friederici AD, Stephan KE, Amunts K. 2007. Effective connectivity of the left BA 44, BA 45, and inferior temporal gyrus during lexical and phonological decisions identified with DCM [Internet]. Hum Brain Mapp. Available from: URL http://dx.doi.org/10.1002/hbm.20512. Accessed 6 January 2009.
Hickok G, Buchsbaum B, Humphries C, Muftuler T. 2003. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. J Cogn Neurosci. 15(5):673–682.
Hickok G, Poeppel D. 2000. Towards a functional neuroanatomy of speech perception. Trends in Cognitive Science. 4(4):131–138.
Hickok G, Poeppel D. 2004. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. Cognition. 92(1–2):67–99.
Hickok G, Poeppel D. 2007. The cortical organization of speech processing. Nat Rev Neurosci. 8(5):393–402.
Indefrey P, Levelt W. 2000. The neural correlates of language production. In: Gazzaniga M, editor. The new cognitive neurosciences. Cambridge (MA): MIT Press. p. 845–865.
Indefrey P, Levelt WJM. 2004. The spatial and temporal signatures of word production components. Cognition. 92(1–2):101–144.
Jusczyk P, Luce P, Charles-Luce J. 1994. Infants' sensitivity to phonotactic patterns in the native language. J Mem Lang. 33:630–645.
Kellman P, van Gelderen P, de Zwart JA, Duyn JH. 2003. Method for functional MRI mapping of nonlinear response. Neuroimage. 19(1):190–199.
Levelt WJM. 1999. Models of word production. Trends Cogn Sci. 3(6):223–232.
Levelt WJM, Roelofs A, Meyer AS. 1999. A theory of lexical access in speech production. Behav Brain Sci. 22(1):1–38.
Levelt WJM, Wheeldon L. 1994. Do speakers have access to a mental syllabary? Cognition. 50(1–3):239–269.
Liberman AM, Mattingly IG. 1985. The motor theory of speech perception revised. Cognition. 21(1):1–36.
Mazaika P, Whitfield-Gabrieli S, Reiss A. 2007. Artifact repair for fMRI data from high motion clinical subjects. Poster presented at: 13th Annual Meeting of the Organization for Human Brain Mapping; Chicago, IL. Hum Brain Mapp Conf.
Molnar-Szakacs I, Iacoboni M, Koski L, Mazziotta JC. 2005. Functional segregation within pars opercularis of the inferior frontal gyrus: evidence from fMRI studies of imitation and action observation. Cereb Cortex. 15(7):986–994.
Oakes TR, Johnstone T, Walsh KSO, Greischar LL, Alexander AL, Fox AS, Davidson RJ. 2005. Comparison of fMRI motion correction software tools. Neuroimage. 28(3):529–543.
Okada K, Hickok G. 2006. Identification of lexical-phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. Neuroreport. 17(12):1293–1296.
Pisoni DB, Nusbaum HC, Luce PA, Slowiaczek LM. 1985. Speech perception, word recognition and the structure of the lexicon. Speech Commun. 4(1–3):75–95.
Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD. 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage. 10(1):15–35.
Riecker A, Mathiak K, Wildgruber D, Erb M, Hertrich I, Grodd W, Ackermann H. 2005. fMRI reveals two distinct cerebral networks subserving speech motor control. Neurology. 64(4):700–706.
Rizzolatti G, Arbib MA. 1998. Language within our grasp. Trends Neurosci. 21(5):188–194.
Rizzolatti G, Craighero L. 2004. The mirror-neuron system. Annu Rev Neurosci. 27:169–192.
Shuster LI, Lemieux SK. 2005. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. Brain Lang. 93(1):20–31.
Vitevitch M, Luce P. 1998. When words compete: levels of processing in perception of spoken words. Psychol Sci. 9(4):325–329.
Vitevitch MS, Luce PA. 2004. A web-based interface to calculate phonotactic probability for words and nonwords in English. Behav Res Methods Instrum Comput. 36(3):481–487.
Vitevitch MS, Luce PA. 2005. Increases in phonotactic probability facilitate spoken nonword repetition. J Mem Lang. 52(2):193–204.
Vitevitch MS, Luce PA, Charles-Luce J, Kemmerer D. 1997. Phonotactics and syllable stress: implications for the processing of spoken nonsense words. Lang Speech. 40(Pt 1):47–62.
Vitevitch MS, Luce PA, Pisoni DB, Auer ET. 1999. Phonotactics, neighborhood activation, and lexical access for spoken words. Brain Lang. 68(1–2):306–311.
Zatorre RJ, Evans AC, Meyer E, Gjedde A. 1992. Lateralization of phonetic and pitch discrimination in speech processing. Science. 256(5058):846–849.
Zatorre RJ, Meyer E, Gjedde A, Evans AC. 1996. PET studies of phonetic processing of speech: review, replication, and reanalysis. Cereb Cortex. 6(1):21–30.