Using Non-Invasive Stimulation of the Undamaged Brain to Guide the Identification of Lesion Sites that Predict Language Outcome After Stroke

by

Diego L. Lorca-Puls

Wellcome Centre for Human Neuroimaging

UCL Queen Square Institute of Neurology

University College London

A dissertation submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Neuropsychology and Cognitive Neuroscience

Supervisors: Professor Cathy J. Price FMedSci

Professor Joseph T. Devlin

Dr. Mohamed L. Seghier

Dr. Thomas M. H. Hope

March 2019

Declaration

I, Diego L. Lorca-Puls, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

_____

Diego L. Lorca-Puls

Publications Statement

The work I have been involved in during my PhD has resulted in the following peer-reviewed publications:

- **Lorca-Puls DL**, Gajardo-Vidal A, Seghier ML, Leff AP, Sethi V, Prejawa S, Hope TMH, Devlin JT, Price CJ. Using transcranial magnetic stimulation of the undamaged brain to identify lesion sites that predict language outcome after stroke. Brain 2017; 140: 1729-42.
- **Lorca-Puls DL**, Gajardo-Vidal A, White J, Seghier ML, Leff AP, Green DW, Crinion JT, Ludersdorfer P, Hope TMH, Bowman H, Price CJ. The impact of sample size on the reproducibility of voxel-based lesion-deficit mappings. Neuropsychologia 2018; 115: 101-11.
- Gajardo-Vidal A, **Lorca-Puls DL**, Crinion JT, White J, Seghier ML, Leff AP, Hope TMH, Ludersdorfer P, Green DW, Bowman H, Price CJ. How distributed processing produces false negatives in voxel-based lesion-deficit analyses. Neuropsychologia 2018; 115: 124-33.
- Gajardo-Vidal A, **Lorca-Puls DL**, Hope TMH, Parker Jones O, Seghier ML, Prejawa S, Crinion JT, Leff AP, Green DW, Price CJ. How right hemisphere damage after stroke can impair speech comprehension. Brain 2018; 141: 3389-404.
- Loughnan R, **Lorca-Puls DL**, Gajardo-Vidal A, Espejo-Videla V, Gillebert CR, Mantini D, Price CJ, Hope TMH. Generalizing post-stroke prognoses from research data to clinical data. Manuscript in preparation.

Particularly important in the context of the current thesis, the bulk of the findings presented in Chapters 3 to 6 has been reported in Lorca-Puls et al. (2017).

Abstract

Disrupting the neural activity in the left anterior supramarginal gyrus (SMG) or opercular part of the left inferior frontal gyrus (pOp) with repetitive transcranial magnetic stimulation (TMS) has been demonstrated to cause a transient slowing of response times during phonologically more than semantically demanding tasks. Likewise, a wealth of functional magnetic resonance imaging (fMRI) studies have shown increased activation in SMG and/or pOp for phonological relative to semantic processing. Here I set out to investigate whether, and how frequently, stroke damage to SMG and/or pOp results in persistent phonological processing impairments in a large sample of 262 right-handed English-speaking adults, who were tested at least 1 year after a left-hemisphere stroke.

In Experiment I, I compared the effect of damage to different parts of SMG and pOp that were defined by regions of interest from either TMS or fMRI studies of phonological processing in neurologically-normal individuals. I found that the incidence of phonological processing impairments was predicted significantly better by the presence or absence of damage to SMG and pOp regions defined by TMS studies than SMG and pOp regions defined by fMRI studies. Moreover, the discriminatory power (for segregating patients with and without phonological abilities) of the TMS sites was not improved further when combined with the fMRI sites. In Experiment II, I adapted the borders of the TMS SMG and pOp regions to include the surrounding grey and white matter where the presence or absence of stroke damage was consistently associated with the presence or absence of phonological processing impairments. The presence or absence of damage to these new TMS-guided regions was able to explain the incidence of phonological impairments better than the original TMS regions, even in a new sample of patients that was entirely independent of the region identification process. In Experiment III, I showed that damage to the TMS-guided regions accounted for the incidence of phonological impairments substantially better than damage to an alternative set of regions derived from voxel-based lesion-deficit mapping techniques that search the whole brain for areas that are most frequently damaged in those with phonological

impairments. However, the best classification accuracy was observed when the analysis took into account a combination of regions from TMS-guided and voxel-based lesion-deficit mapping approaches. In Experiment IV, I investigated the nature of the functional impairment caused by SMG or pOp lesions and found that damage to either region impaired covert and overt phonological processing abilities more than semantic processing abilities, as predicted by prior TMS and fMRI studies of neurologically-normal subjects. Finally, the behavioural effects of damage were remarkably similar (i.e. no statistically significant differences) for both TMS-guided sites (i.e. pOp and SMG).

In conclusion, the fact that damage to the TMS-guided SMG and pOp regions impaired phonological processing abilities years after stroke onset, suggests that these regions are critical for accurate phonological processing (both overt and covert) and that other brain areas are not typically able to fully compensate for the contribution that these regions make to language processing. More broadly, the results illustrate how non-invasive stimulation of the undamaged brain can be used to guide the identification of regions where brain damage is likely to cause persistent behavioural effects. By combining these regions of interest with those derived from other lesion-deficit mapping approaches, I was not only able to explain the presence, but also the absence, of phonological processing impairments in a large cohort of patients.

Impact Statement

More than 13 million people worldwide suffer a stroke every year, and around a third of all incident cases are left with residual language disorders commonly referred to as aphasia. Post-stroke aphasia is one of the most devastating behavioural consequences of stroke leading to an increased risk of depression, social isolation and inability to work. However, clinicians face outstanding limitations when generating prognoses about language outcome and recovery after stroke because of our current state of knowledge in the field. This situation is further aggravated by the fact that stroke survivors and their relatives typically want to know what will happen to them and the likelihood of making a good recovery over time. In addition, post-stroke aphasia poses a substantial burden on social security and public health systems due to the long-term needs - both in terms of rehabilitation and care - of the affected patients.

The research work reported in my thesis is intended to help bridge this gap by illustrating how previous findings from non-invasive brain stimulation studies of neurologically-normal individuals can be used to guide the search for brain regions where stroke damage consistently disrupts language abilities. In other words, I focused on investigating ways to improve our ability to predict language outcome and recovery after stroke by studying the incidence of speech-sound processing impairments at the individual patient level rather than on testing the statistical significance of group-level effects (that may be driven by a subset of patients). This work may thus be relevant to all those in the lesion-behaviour mapping community who aim to understand the neurobiology of language without sacrificing the clinical translatability of their research outputs. For example, by using this approach, I have identified two specific lesion sites in frontal and parietal brain areas that very consistently result in speech-sound processing impairments during the first 5 years post-stroke. With further validation, these findings may in the future be able to inform clinical practice including prognosis, goal setting and therapy planning. Finally, the methodological approach presented in my

thesis could easily be extended to answer pressing questions in other behavioural domains such as memory and movement.

Acknowledgements

First and foremost, I would like to thank my primary supervisor, Professor Cathy Price, whose guidance, support and mentorship have been absolutely fundamental for the completion of this PhD thesis. I have learnt so many invaluable lessons from you over these past few years that I would not be able to translate them all into words. However, there is one in particular I would like to mention here: you have taught me how to carry out research to the highest standards of scientific and ethical rigour. And for that, I will always be indebted to you.

In addition to the critical role my primary supervisor has played throughout my PhD, I am grateful to my secondary supervisors, Professor Joseph Devlin, Dr. Mohamed Seghier and Dr. Thomas Hope, who always made themselves available to provide me with guidance and support; particularly at those (multiple) points of the PhD where my research work proved to be more difficult and burdensome than I anticipated.

This thesis is wholeheartedly dedicated to my beloved parents (Alejandro and Liliana) and wife (Andrea). Dad and Mum, you have always been an endless source of inspiration for me and I will continue to strive to live up to the example set by my two most important role models. Andrea, you are everything to me, the greatest companion life could have made me cross paths with. All the reflections, experiences and emotions we have shared throughout this incredible adventure are my most precious treasure. Holding your hand is all that is needed to remind me I belong somewhere: I belong with you.

To both of my sisters (Muriel and Alejandra), you are my family and I deeply appreciate the bond that unite us. I could not be more grateful to you for bringing into existence my dearests Antonia, Catalina and Mateo. Being in their presence connects me with an aspect of the self that I had mistakenly thought lost and it is with great excitement that I look forward to the next opportunity where we will all be together again. I love you.

To my parents- (Andres and Grette) and sisters-in-law (Grette and Ingrid), my deepest gratitude for always making me feel welcome and at home. To the lovely Amparo and Rafaela, thank you just for being the beautiful little persons you are.

I would also like to thank all the amazing people who, since I firstly arrived in London, have helped me to broaden my horizons and step out of my comfort zone through their friendship: Sangho, Geehyun, Sharif, Claire, Ana (Gonzalez), Natalie, Marion, Julie, Philipp, Berk, Conny, Eugenia, Patrick, Oscar, Daniela, Ellie, Sharifa, Iris, Jane and Phil.

I am grateful to my wonderful colleagues in the language group (past and current members) for creating such a friendly and supportive environment, which has made my work so much easier and so rewarding: Louise, Johanna, Zula, Rachel, Hinal, Oiwi, Ana (Sanjuán), Elnas, Andre, Deborah, Suz, Sue, David (Green), Howard, Sharon, Justyna, Letitia, Storm, Zoe, Emilia, Valeria, Kate, Hayley, Shumi, Sophie and Megan (Docksey).

My special acknowledgements go to all those people (admin, imaging, IT, methods, physics, PhD students, postdocs and PIs) who have contributed with their work to make the FIL one of the most inspiring and best research centres around the world. I am especially grateful to: Kamlyn, Monica, Marcia, Elaine, Shiv, Maddy, Marina, Nathalie, Peter, David (Bradbury), Rachael, Chris, Liam, Jorge, Cassie and Megan (Creasey). I am also indebted to the many people who have made my office the best office (by far) in the FIL: Spas, Eleanor, Ryszard, Becky, Phillip, Berk, Conny, Daniel, Tim and Anna.

I am extremely grateful to the Chilean National Commission for Scientific and Technological Research (CONICYT) for funding my PhD through its scholarship scheme for graduate studies abroad (CONICYT BECAS-CHILE 72140131).

Last but not least, thank you to all the stroke survivors (and their relatives) who accepted to take part in the PLORAS project. Without their generosity and patience this thesis would not have been possible. I hope my research work will eventually have a

positive impact on those whose life has been dramatically affected by the occurrence of a stroke. That is certainly one of the key reasons why I decided to leave the clinic and pursue a career as a neuroscientist. My commitment to you is to conduct much needed clinically relevant research.

Table of Contents

CHAPTER 1: General Introduction 21

1.1 The burden of stroke 22

1.2 Aphasia: a devastating behavioural consequence of stroke 23

1.3 A syndrome-based aphasia classification system 24

1.4 Inconsistencies in the lesion sites associated with major aphasic syndromes 27

1.5 Towards a better understanding of the relationship between lesion site and language impairments after stroke 29

1.6 A psycholinguistic perspective on phonology and phonological processing 30

1.6.1 Phonology 30

1.6.2 Phonological processing 31

1.6.3 Phonological working memory 34

1.7 Neurobiology of phonological processing 36

1.7.1 Neurobiology of phonological processing in neurologically-normal individuals 36

1.7.2 Neurobiology of phonological processing in clinical populations 40

1.7.3 Phonological working memory as an emergent property of the language system 42

1.7.4 Neurocomputational models of speech processing 44

1.8 Lesion-deficit mapping studies 47

1.8.1 Interpreting univariate voxel-based lesion-deficit mappings 48

1.8.2 Interpreting multivariate machine-learning-based lesion-deficit mappings 51

1.9 Thesis overview 55

CHAPTER 2: General Materials and Methods 56

2.1 The PLORAS database 57

2.1.1 My contribution to the PLORAS database 59

2.2 The behavioural assessment tools 60

2.3 Phonological and semantic decisions 70

2.4 Basic principles of MRI physics 73

2.5 Structural MRI data acquisition, pre-processing and lesion identification 78

2.6 Neurophysiological basis of the fMRI BOLD signal 84

2.7 Biophysical foundations of TMS 86

2.8 Prediction accuracy 89

CHAPTER 3 (Experiment I): Predicting Phonological Processing Impairments from Stroke Damage to Regions Identified by TMS and fMRI Studies of Normal Phonological Processing 91

3.1 Summary 92

3.2 Introduction 93

3.2.1 Research question 96

3.3 Materials and Methods 97

3.3.1 Participants 97

3.3.2 Assessing phonological processing abilities 102

3.3.3 Cross-validating the definition of phonological processing impairments 103

3.3.4 Regions of interest from prior TMS and fMRI studies 105

3.3.5 Determining the threshold for critical damage 110

3.4 Results 114

3.4.1 Cross-validating the definition of phonological processing impairments 114

3.4.2 Critical damage thresholds 115

3.4.3 Classification accuracy for TMS regions 116

3.4.4 Classification accuracy for fMRI regions 117

3.4.5 The real extent of the lesions in patients with damage to the TMS regions and phonological impairments 118

3.5 Discussion 119

3.5.1 The predictive value of the TMS regions was surprisingly high 120

3.5.2 The TMS regions are persistently necessary for accurate phonological processing 121

3.5.3 The lower predictive value of the fMRI compared to TMS regions was not surprising 121

3.5.4 Explaining inter-patient variability in the effect of damage to the same regions 123

3.5.5 Limitations 124

3.6 Conclusions 125

CHAPTER 4 (Experiment II): Improving the Prediction Accuracy of the TMS Regions 127

4.1 Summary 128

4.2 Introduction 129

4.2.1 Research question 132

4.3 Materials and Methods 132

4.3.1 Participants 132

4.3.2 Defining phonological processing impairments 134

4.3.3 Generating new TMS-guided and fMRI-guided regions 135

4.3.4 Determining the threshold for critical damage 136

4.4 Results 137

4.4.1 The TMS-guided regions 137

4.4.2 The fMRI-guided regions 139

4.4.3 Critical damage thresholds 140

4.4.5 Classification accuracy for TMS-guided regions in Sample 1 141

4.4.6 Comparing the classification accuracy for TMS and TMS-guided regions . 143

4.4.6.1 Sample 1 143

4.4.6.2 Sample 3 145

4.4.7 Comparing the classification accuracy for fMRI-guided and TMS-guided regions 146

4.4.7.1 Sample 1 146

4.4.7.2 Sample 3 147

4.4.8 Inconsistency in the effect of TMS-guided pOp lesions in Sample 3 147

4.5 Discussion 149

4.5.1 The effect of damage to the TMS-guided regions on phonological processing 150

4.5.2 The depth and extent of the TMS effect on phonological processing 151

4.5.3 Performance was not worse following damage to both regions than SMG or pOp alone 153

4.5.4 Focusing on the consistency of lesion effects rather than their statistical significance 153

4.5.5 Inter-patient variability in the effect of damage to the TMS-guided regions 154
4.5.6 Limitations ... 155
4.6 Conclusions ... 157
CHAPTER 5 (Experiment III): Comparing the Prediction Accuracy of the TMS-guided Regions with that of a Set of Regions Derived from Voxel-based Lesion-Deficit Mapping Analyses ... 159
5.1 Summary ... 160
5.2 Introduction ... 162
5.2.1 Research question ... 167
5.3 Materials and Methods ... 167
5.3.1 Participants ... 167
5.3.2 Regions of interest from univariate lesion-deficit mapping analyses ... 168
5.3.3 Determining the threshold for critical damage ... 170
5.3.4 Evaluating the combination of regions that maximised classification accuracy ... 170
5.4 Results ... 170
5.4.1 Regions of interest from univariate lesion-deficit mapping analyses ... 170
5.4.2 Critical damage thresholds ... 173
5.4.3 Classification accuracy for Sample 1 ... 174
5.4.4 Classification accuracy for Sample 3 ... 175
5.4.5 Logistic regression analysis ... 177
5.5 Discussion ... 178
5.5.1 The lesion sites that best explain the incidence of phonological impairments ... 178
5.5.2 The seed TMS regions increased classification accuracy ... 179
5.5.3 The voxel-based statistical analyses failed to detect the TMS-guided regions ... 180
5.5.4 The voxel-based statistical analyses identified similar regions ... 181
5.5.5 Limitations ... 183
5.6 Conclusions ... 184
CHAPTER 6 (Experiment IV): Investigating the Functional Role of the TMS-guided SMG and pOp Regions ... 186

6.1 Summary 187
6.2 Introduction 188
6.2.1 Research question 190
6.2.2 Rationale 190
6.3 Materials and Methods 191
6.3.1 Participants 191
6.3.2 Task analysis 191
6.3.3 Comparing the effect of damage to the TMS-guided regions on covert versus overt phonological processing 192
6.3.4 Comparing the effect of damage to the TMS-guided regions on phonological versus semantic processing abilities 193
6.3.5 Testing for differences in the effect of damage to TMS-guided SMG versus pOp 194
6.4 Results 194
6.4.1 Comparing the effect of damage to the TMS guided-regions on covert versus overt phonological processing 194
6.4.2 Comparing the effect of damage to the TMS-guided regions on phonological versus semantic processing abilities 194
6.4.3 Testing for differences in the effect of damage to TMS-guided SMG versus pOp 195
6.5 Discussion 196
6.5.1 Covert versus overt phonological processing 198
6.5.2 Phonology versus semantics 199
6.5.3 TMS-guided SMG versus TMS-guided pOp 199
6.5.4 Limitations 200
6.6 Conclusions 202
CHAPTER 7: General Discussion 203
7.1 The predictive power of the TMS regions is better than that of the fMRI regions 204
7.2 The lesion sites underlying persistent phonological impairments involve a combination of grey and white matter 206
7.2.1 Inconsistency in the effect of damage to the TMS-guided pOp region 206

7.3 Phonological processing abilities after stroke are best explained by considering the presence or absence of damage to regions identified by TMS-guided and voxel-based lesion-deficit mapping analyses 208
7.4 The TMS-guided regions are involved in covert phonological processing 210
7.5 An 8-step procedure for identifying regions that predict outcome after stroke 210
7.6 Conclusions 212
References 213
Appendix 1 (Chapter 2) 245
1 Visual phonological decision task 245
2 Visual semantic decision task 247

List of Figures

Figure 1.1: Dual-stream models of the language network in the human brain. 38
Figure 2.1: The PLORAS approach. 60
Figure 2.2: MRI scanner. 74
Figure 2.3: MRI signal generation process. 75
Figure 2.4: MRI signal location and encoding. 77
Figure 2.5: Automated lesion identification 83
Figure 2.6: The fMRI BOLD signal. 85
Figure 2.7: Biophysical foundations of TMS. 87
Figure 2.8: TMS protocols. 88
Figure 3.1: Lesion overlap map for Sample 1 99
Figure 3.2: Lesion overlap map for Sample 2. 100
Figure 3.3: Comparison of the distribution of lesions between Sample 1 and Sample 2. 102
Figure 3.4: Task analysis. 104
Figure 3.5: TMS regions of interest. 106
Figure 3.6: fMRI SMG regions of interest. 109
Figure 3.7: fMRI pOp regions of interest. 110
Figure 3.8: Lesion overlap map for Subsets 1A and 1B. 113
Figure 3.9: Strength of the association between TMS and CAT scores 115
Figure 3.10: Smallest lesion sites. 119
Figure 4.1: Lesion overlap map for Sample 3. 133
Figure 4.2: Comparison of the distribution of lesions between Sample 1 and Sample 3. 134
Figure 4.3: TMS-guided pOp region. 138
Figure 4.4: TMS-guided SMG region. 138
Figure 4.5: fMRI-guided pOp region. 139
Figure 4.6: fMRI-guided SMG region. 140
Figure 4.7: Classification accuracy for TMS and TMS-guided regions. 144
Figure 5.1: Univariate lesion-deficit mapping regions. 172
Figure 5.2: Co-occurring patterns of damage. 173
Figure 6.1: Task analysis. 192
Figure 6.2: Incidence and severity of impairments. 197

List of Tables

Table 1.1: The Boston syndrome-based aphasia classification system. 26
Table 2.1: 2×2 contingency table. 89
Table 3.1: Summary of demographic and clinical data for Sample 1. 99
Table 3.2: Summary of demographic and clinical data for Sample 2. 100
Table 3.3: Comparison of CAT scores between Sample 1 and Sample 2. 101
Table 3.4: Summary of demographic and clinical data for Subsets 1A and 1B. 113
Table 3.5: Critical damage thresholds for TMS and fMRI regions. 116
Table 3.6: Classification accuracy for TMS and fMRI regions in Sample 1. 117
Table 3.7: Overlap between spherical TMS and fMRI regions. 118
Table 4.1: Summary of demographic and clinical data for Sample 3. 133
Table 4.2: Comparison of CAT scores between Sample 1 and Sample 3. 135
Table 4.3: Overlap between TMS, TMS-guided and fMRI-guided regions. 140
Table 4.4: Critical damage thresholds for TMS-guided and fMRI-guided regions. 141
Table 4.5: Summary of demographic and clinical data for lesion-size-matched groups. 142
Table 4.6: Comparison of the incidence and severity of phonological processing impairments. 143
Table 4.7: Classification accuracy for TMS, TMS-guided and fMRI-guided regions in Sample 1. 144
Table 4.8: Lesion categorisation. 145
Table 4.9: Classification accuracy for TMS, TMS-guided and fMRI-guided regions in Sample 3. 146
Table 4.10: Classification accuracy for TMS-guided and fMRI-guided pOp regions. 147
Table 5.1: Overlap between TMS-guided and univariate lesion-deficit mapping regions. 173
Table 5.2: Critical damage thresholds for univariate lesion-deficit mapping regions. 174
Table 5.3: Classification accuracy for TMS-guided and univariate lesion-deficit mapping regions in Sample 1. 175
Table 5.4: Classification accuracy for TMS-guided and univariate lesion-deficit mapping regions in Sample 3. 176
Table 5.5: Incidence of phonological processing impairments according to the damage status of the best combination of regions. 177
Table 6.1: Detrimental effect of damage to the TMS-guided regions on covert and overt phonological processing. 195
Table 6.2: Preferential involvement of the TMS-guided regions in phonological versus semantic processing. 196

Abbreviations

a	Anterior
AF	Arcuate fasciculus
aka	Also known as
BOLD	Blood-oxygenation-level-dependent
CAT	Comprehensive aphasia test
TBS	Theta-burst stimulation
ECoG	Electrocorticography
EmC	Extreme capsule
fMRI	Functional magnetic resonance imaging
FWE	Family-wise error
IFG	Inferior frontal gyrus
LFP	Local field potential
LOM	Lesion overlap map
MNI	Montreal Neurological Institute
MRI	Magnetic resonance imaging
p	Posterior
PLORAS	Predicting language outcome and recovery after stroke
PMC	Premotor cortex
pOp	Pars opercularis of the inferior frontal gyrus
PostCG	Postcentral gyrus
PreCG	Precentral gyrus
RF	Radiofrequency
ROI	Region of interest
SLF	Superior longitudinal fasciculus
SMG	Supramarginal gyrus
SPM	Statistical parametric mapping

STG	Superior temporal gyrus
STS	Superior temporal sulcus
SVR	Support vector regression
TE	Echo time
TMS	Transcranial magnetic stimulation
TR	Repetition time
v	Ventral
VBM	Voxel-based morphometry
VLSM	Voxel-based lesion-symptom mapping

CHAPTER 1:

General Introduction

The need for better biomarkers of stroke outcome and recovery has been advocated by a taskforce of international researchers (Bernhardt et al., 2017; Boyd et al., 2017). This is because being able to accurately predict long-term outcome and recovery after stroke is anticipated to: (i) help alleviate patients' distress and caregivers' burden; (ii) guide clinicians in goal setting and therapy planning; and (iii) inform the stratification of patients into different groups in clinical trials (based on the predicted potential for recovery; e.g., Stinear et al., 2012; Hope et al., 2013). In that sense, a large number of studies have shown that lesion site is a major determinant of outcome and recovery after stroke (e.g., Marchina et al., 2011; Fridriksson et al., 2013; Wang et al., 2013; Forkel et al., 2014; Hillis et al., 2018).

The overarching goal of this thesis is to investigate whether findings from prior transcranial magnetic stimulation and functional magnetic resonance imaging studies of the neurologically-normal brain can be used to guide the identification of brain regions where stroke damage consistently predicts the incidence of language processing impairments. Below, I contextualise the pertinence of this work in clearly labelled sections that distil all the necessary background information for the reader. The first half delves into a range of issues including (i) the burden of acquired language disorders (i.e. aphasia) post-stroke, (ii) the limited prognostic value of the classic aphasia classification system that relies on symptom clusters, and (iii) the need for better biomarkers of long-term language outcome and recovery after stroke. The second half, in turn, focuses on (iv) analysing the language function that will be examined to demonstrate the utility of the proposed methodological procedure from psycholinguistic and neurobiological angles, and (v) illustrating how recent and mainstream lesion-deficit mapping techniques have not yet delivered a satisfactory method for generating accurate outcome predictions at the individual patient level.

1.1 The burden of stroke

Stroke is the largest contributor to the burden of neurological disorders worldwide and the second leading cause of mortality as well as disability across a wide spectrum

of diseases and injuries (Feigin et al., 2017; Hay et al., 2017; Naghavi et al., 2017). Globally, it has been recently estimated that (i) there are about 79.6 million stroke survivors alive today, (ii) nearly 5.5 million people die from stroke on a yearly basis, (iii) 116.4 million disability-adjusted life-years are due to stroke and (iv) more than 13 million people are affected by first-ever stroke each year (Hay et al., 2017; Naghavi et al., 2017; Vos et al., 2017). Over all reported cases, approximately two-thirds correspond to ischaemic strokes and one-third to haemorrhagic strokes (Vos et al., 2017), with 31% of them occurring in young to middle-aged adults (i.e. 20-64 years of age) (Krishnamurthi et al., 2015). The associated global burden and costs of stroke are expected to continue to rise in the near future due to population growth and aging, longer life expectancy, improved stroke care and higher prevalence of modifiable risk factors (Carter et al., 2007; Roth et al., 2015; Feigin et al., 2016a, b; O'Donnell et al., 2016).

1.2 Aphasia: a devastating behavioural consequence of stroke

One of the most devastating and disabling behavioural consequences of stroke is aphasia, which can be defined as an acquired language disorder that typically disrupts the ability to speak, comprehend, write and/or read (Berthier, 2005; Berthier and Pulvermüller, 2011; Tippett and Hillis, 2017). The median frequency of post-stroke aphasia across multiple epidemiological studies has been calculated to be around 30% in acute settings (i.e. in-patients) and 34% in chronic settings (i.e. out-patients). Although there is substantial inter-study variation in these figures (range = 20%-41% and 25%-52% for acute and chronic settings, respectively; Flowers et al., 2016), the evidence allows us to conclude that approximately 4 million people worldwide are affected by post-stroke aphasia every year (Engelter et al., 2006; Tsouli et al., 2009; Croquelois and Bogousslavsky, 2011).

The magnitude of the challenge posed by post-stroke aphasia is complicated by its positive relationship with age (Engelter et al., 2006; Tsouli et al., 2009; Croquelois and Bogousslavsky, 2011; Ellis and Urban, 2016), particularly when considering that the latest projections from the United Nations indicate that by 2030 roughly 1 billion people

will be 65 years old or over (United Nations, 2017). Furthermore, compared to those without aphasia, stroke patients with aphasia are at greater risk of long hospital stays, anxiety, depression, dependence and permanent unemployment (Tsouli et al., 2009; Graham et al., 2011; Tanaka et al., 2014; Shehata et al., 2015; Boehme et al., 2016; Morris et al., 2017; Pike et al., 2017). These factors in combination with other cognitive and non-cognitive comorbidities lead patients with post-stroke aphasia to experience social isolation and long-term disability (Dickey et al., 2010; Hilari and Northcott, 2017; Lazar and Boehme, 2017; Wray and Clarke, 2017), increasing societal costs due to the high health care expenditures associated with meeting the long-term needs of those who suffer from persistent post-stroke aphasia (Ellis et al., 2012; Boehme et al., 2016).

The issues summarised above therefore speak directly to the urgency of improving the treatment of aphasia and the need to know who will benefit most from it. Indeed, finding the best ways to help people recover from aphasia has been recognised as one of the top ten research priorities relating to life after stroke (Pollock et al., 2014; Franklin et al., 2018). Currently, clinicians have very limited knowledge to deal with patients' and relatives' expectations and enquiries about prognosis (Worrall et al., 2011; Howe et al., 2012) since the main determinants of language outcome after stroke as well as the neural mechanisms that support recovery from aphasia are still poorly understood (Kiran, 2012), as reflected by the fact that gold standard pharmacotherapies are yet to be developed (Berthier et al., 2011; Llano and Small, 2016). Encouragingly, however, accumulating evidence shows that high-intensity/high-dose speech and language therapy may be beneficial, under specific circumstances (e.g., Aguilar et al., 2018), for patients with post-stroke aphasia even when delivered during the chronic phase (Allen et al., 2012; Brady et al., 2016; Breitenstein et al., 2017; Stahl et al., 2018).

1.3 A syndrome-based aphasia classification system

Clinico-anatomical correlations in aphasiology have been guided by the seminal observations from Paul Broca (1861a, b) and Carl Wernicke (1874) for over 100 years (Dronkers et al., 2017). More than a century ago Broca described a patient with impaired

speech production who had a lesion in the posterior part of the left inferior frontal gyrus. Wernicke, in turn, reported that lesions to the posterior aspect of the left superior temporal gyrus caused language comprehension deficits. In honour of their original contributions, the posterior part of the left inferior frontal gyrus and the posterior part of the left superior temporal gyrus were named Broca's area and Wernicke's area, respectively (Lazar and Mohr, 2011). Subsequently, Ludwig Lichtheim (1885) complemented and extended Wernicke's work by introducing the existence of two other types of aphasia that later became known as transcortical motor aphasia and transcortical sensory aphasia (Heilman, 2006; Catani and Mesulam, 2008; Cauquil-Michon et al., 2011). However, Lichtheim himself did not ascribe any specific anatomical location to these new aphasic syndromes.

Building upon these neurological findings, Norman Geschwind (1970) drew attention to the importance of the arcuate fasciculus in connecting Broca's and Wernicke's areas. He reasserted (following Carl Wernicke's original claims) that lesions severing this white matter tract result in a disconnection syndrome called conduction aphasia whose pathognomonic symptom is impaired repetition (Geschwind, 1970). Furthermore, Geschwind and his co-workers at the Boston University Aphasia Research Centre synthetized all the evidence accumulated thus far into a syndrome-based aphasia classification system with coarse-grained lesion correlates (Geschwind, 1965a, b; Goodglass et al., 2001). A distinctive feature of the Boston school of aphasiology is that it places a particular emphasis on distinguishing non-fluent from fluent forms of aphasia (Geschwind, 1971); see Table 1.1 for details. Crucially, although there are alternative schools of thought in aphasiology (for an example, see Ardila, 2010), the Boston syndrome-based aphasia classification system remains the most influential one to date.

Table 1.1: The Boston syndrome-based aphasia classification system.

Aphasic syndrome	Description
Non-fluent forms of aphasia (Damasio and Geschwind, 1984; Hillis, 2007; Tippett and Hillis, 2016):	
Broca's	Broca's aphasia is typically associated with damage to Brodmann areas (BA) 44 and 45 and characterised by non-fluent agrammatic conversational speech, impaired repetition and relatively preserved language comprehension.
Transcortical motor	Transcortical motor aphasia is typically associated with prefontal damage located anteriorly or superiorly to Broca's area and defined by non-fluent speech output and relatively well-preserved comprehension and repetition (unlike Broca's aphasia).
Mixed transcortical	Mixed transcortical aphasia is typically associated with damage to brain regions neighbouring Broca's and Wernicke's areas and resembles global aphasia in that language production and comprehension are compromised but with relatively spared repetition and echolalia.
Global	Global aphasia is the most devastating form of aphasia typically resulting from large lesions involving extensive portions of the left perisylvian territory and characterised by severely impaired language production and comprehension in addition to poor repetition skills.
Fluent forms of aphasia (Damasio and Geschwind, 1984; Hillis, 2007; Tippett and Hillis, 2016):	
Anomic	Anomic aphasia is the mildest form of aphasia and refers to a syndrome characterised by noticeable word finding difficulties (i.e. anomia) in the context of otherwise relatively intact language abilities.
Conduction	Conduction aphasia is typically associated with lesions to the left arcuate fasciculus and characterised by relatively fluent spontaneous speech, phonemic paraphasias, impaired repetition and relatively spared language comprehension.
Wernicke's	Wernicke's aphasia is typically associated with lesions centred on the posterior end of BA 22 and defined by fluent but paraphasic speech production with little or no content, impaired language comprehension and repetition.
Transcortical sensory	Transcortical sensory aphasia is typically associated with parieto-temporal damage outside the putative left perisylvian language network and shares many of the features of Wernicke's aphasia but with relatively spared repetition.

1.4 Inconsistencies in the lesion sites associated with major aphasic syndromes

Despite the significant contribution that the Boston syndrome-based aphasia classification scheme has made to facilitating the communication of complex behavioural phenomena in the clinical setting, its utility for advancing aphasia research is at the very least controversial (Caramazza, 1984; Caplan, 1993; Marshall, 2010); particularly if we are to understand how language outcome and recovery after stroke are related to lesion site (e.g., Price et al., 2010a). For instance, the traditional association between non-fluent aphasia with anterior (or pre-Rolandic) lesions and fluent aphasia with posterior (or post-Rolandic) lesions (Benson, 1967) has been challenged on the basis of research findings that have shown that "exceptions" to these lesion localisation rules are observed with far greater frequency than expected (Basso et al., 1985; Willmes and Poeck, 1993).

In the same vein, MRI examination of the brains from the two historic cases reported by Broca has shown that the stroke lesions were not restricted to Broca's area and included other brain structures such as deep white matter tracts (Dronkers et al., 2007). This is in keeping with prior reports that focal stroke damage to Broca's area only results in a transient language disorder that initially manifests as mutism and that far more extensive damage incorporating the grey and white matter surrounding Broca's area is needed to observe the full constellation of symptoms referred to as Broca's aphasia (Mohr et al., 1978; see also Alexander et al., 1990). Critically, a recent study found that chronic Broca's aphasia appears to originate from a combination of Broca's and Wernicke's areas damage (Fridriksson et al., 2015). Moreover, it has been demonstrated that Broca's aphasia can occur in the absence of evident damage to Broca's area (Marie and Moutier, 1906; Fridriksson et al., 2007).

Similar findings have also been recorded for Wernicke's aphasia where (for a critical appraisal of the evidence, see Binder, 2015, 2017): (i) selective damage to Wernicke's area has been associated with good and rapid recovery of language function (Selnes et al., 1984; Yagata et al., 2017); (ii) persistent Wernicke's aphasia has been

shown to result from lesions that involve the left supramarginal and angular gyri in addition to Wernicke's area (Kertesz et al., 1993); (iii) the presence of Wernicke's aphasia has been documented in the absence of Wernicke's area damage (Roh et al., 2009); and (iv) structural abnormality in Wernicke's area has not been found to be a significant contributor to language comprehension impairments (i.e. the hallmark symptom of Wernicke's aphasia) (Dronkers et al., 2004; Mesulam et al., 2015). Furthermore, even among language neuroscientists there is substantial disagreement regarding the precise anatomical location of Broca's and Wernicke's areas, which has led some authors to argue that these terms should be completely abandoned (Tremblay and Dick, 2016).

Conduction aphasia as well as the remaining aphasia types have also been subject to criticism on related grounds (e.g., Basso et al., 1985; Vignolo et al., 1986; Willmes and Poeck, 1993; Selnes et al., 2002; Bernal and Ardila, 2009; Fridriksson et al., 2010; Croquelois and Bogousslavsky, 2011; Epstein-Peterson et al., 2012; Kasselimis et al., 2017). In addition, contrary to what might be assumed, the various forms of aphasia are not stable entities but change in unpredictable ways during the recovery process (Willmes and Poeck, 1993; Laska et al., 2001; Pedersen et al., 2004). Likewise, it has become increasingly clear that the clusters of symptoms previously described arise from far more complex damage patterns than originally thought (Henseler et al., 2014; Yourganov et al., 2015; Fridriksson et al., 2018). Crucially, a substantial number of patients with post-stroke aphasia cannot be classified because their behavioural profiles do not fit well with the definition of any of the major aphasic syndromes (Crary et al., 1992; Hoffmann and Chen, 2013; Kasselimis et al., 2017). Finally, recent advances in neuroimaging techniques in combination with the study of other clinical populations such as patients with primary progressive aphasia have helped to elucidate that areas outside the putative left-lateralised language network are involved in language processing (Friederici, 2011; Price, 2012; Dick et al., 2014; Mesulam et al., 2014; Lambon Ralph et al., 2017). The evidence presented above therefore calls for

substantial revision of the classic neurological model of aphasia (Charidimou et al., 2014; Chang et al., 2015; Dronkers et al., 2017).

1.5 Towards a better understanding of the relationship between lesion site and language impairments after stroke

One of the challenges of classifying patients by their aphasic symptomatology is that any particular aphasic symptom (e.g., anomia) can originate from disruption to different levels of processing given the multifaceted nature of language tasks (DeLeon et al., 2007; Schwartz, 2014; Blumstein, 2016). For instance, in order to name the picture of a "bed" (i.e. a widely used language task) a person has at least to be able to: (1) recognise the image at hand; (2) access the meaning of the object; (3) retrieve the word form; and (4) control and move the articulators (e.g., tongue, lips, etc.) to pronounce the word "bed". An inability to name objects could result from multiple dissociable functional impairments depending on the level/levels of processing that has/have been affected. In addition, current research into the neurobiology of language has demonstrated that even the simplest language function is supported by a network of regions (i.e. a neural system) (Mesulam, 1990; Vigneau et al., 2006; Price, 2010, 2012; Fuertinger et al., 2015). Thus, it is not surprising that the current classification system fails to capture the complexity of the structure-function relationships underlying the deficits observed in patients with post-stroke aphasia.

A more promising approach may entail shifting the focus from the study of syndromes or symptoms to investigations of very specific functional impairments that can be more easily mapped to discrete levels of psycholinguistic processing (Caramazza, 1984; Caplan, 1993; Blumstein, 2016). For instance, El Hachioui et al. (2013a) tested a sample of 147 patients with post-stroke aphasia at multiple time points during the first year after onset. Patients' language abilities were described in terms of behavioural performance on phonologically, semantically and syntactically demanding tasks. The authors showed that performance on phonological tasks within the first week since stroke was the strongest predictor of 1-year language outcome, even after

adjusting for the effect of other behavioural, demographic and clinical factors such as semantic and syntactic performance at baseline. Furthermore, the authors showed that the phonological, semantic and syntactic processing abilities of the patients followed separate recovery trajectories with phonology improving up to 3 months after stroke (El Hachioui et al., 2013b). This suggests that psycholinguistic classification might be relevant for predicting recovery; specially in the language domain where it is increasingly recognised that behavioural changes (i.e. improvement, maintenance or decline) continue to occur even years after stroke onset (e.g., Holland et al., 2017; Hope et al., 2017) contrary to what has traditionally been held (e.g., Demeurisse et al., 1980; Lendrem and Lincoln, 1985; Pedersen et al., 1995; Laska et al., 2001).

Following the interest in psycholinguistic categorisation of aphasia, my PhD experiments investigate (i) phonological processing impairments after stroke; (ii) how these impairments are related to lesion site; and (iii) whether their incidence can be predicted by lesion site. The reasons I have focused on phonological processing impairments are that: (a) there is a strong relationship between acute phonological processing performance and long-term post-stroke aphasia severity (e.g., Blom-Smink et al., 2017); and (b) phonological processing impairments occur more frequently after stroke than deficits affecting other linguistic levels (e.g., El Hachioui et al., 2012). In the next two sections, I will therefore (a) give a brief overview of how phonology and phonological processing are conceived from a psycholinguistic perspective and (b) summarise the main findings of studies looking into the neurobiology of phonological processing.

1.6 A psycholinguistic perspective on phonology and phonological processing

1.6.1 Phonology

Phonology is concerned with the study of the mental representations of the sound system of human languages, including the underlying principles commanding the composition and combination of speech sounds (Goldrick and Rapp, 2007; Baković,

2014; Idsardi and Monahan, 2016; Nathan, 2017). When learning to speak, we acquire very detailed knowledge about the sound structure of a word (i.e. word-form) and will find it relatively easy to recognize and pronounce unfamiliar speech composed of phonological features we are acquainted with (i.e. the mental representation of the sound structure and meaning of a word can be dissociated; Goldrick and Rapp, 2007).

In psycholinguistics, phonological analysis involves the detailed examination of the regularities of the sound structure of words and the combinatorial rules of speech sounds and their sub-parts (Goldrick and Rapp, 2007; Baković, 2014; Idsardi and Monahan, 2016; Nathan, 2017). Of note, speech sounds are thought to be mentally represented as abstract (i.e. context-independent) units that code the distinctive features of the articulatory gestures/auditory consequences associated with the realization of a particular speech sound when contrasted with that of another (Kazanina et al., 2018). For instance, [p] and [b] are both bilabial stop consonants, meaning that their articulation involves the complete constriction (i.e. stop) of the lips (i.e. bilabial). However, [p] can be distinguished from [b] because [p] does not require the vibration of the vocal folds (i.e. is a voiceless bilabial stop consonant), whereas [b] does (i.e. is a voiced bilabial stop consonant). It therefore follows that "voicing" is a distinctive feature of phonology. Furthermore, phonological theory posits that different levels of abstraction are necessary to mentally represent the sound patterns of a language. For example, phonemes (which link temporally organised groups of distinctive features together) and syllables (specifying the metrical structure of utterances) are the most widely accepted psycholinguistic units (Baković, 2014; Goldrick, 2014). Importantly, these phonological representations are likely to be grounded in the cortical motor and sensory (somatosensation and audition) systems implicated in the production and perception of speech (Hickok, 2014).

1.6.2 Phonological processing

Phonological processing can, in this context, be defined as the mental retrieval and manipulation of phonological representations for the perception and production of

speech (Vitevitch and Luce, 1998, 1999; Goldrick and Rapp, 2007; Robson et al., 2012; Goldrick, 2014; Binder, 2016; Dial and Martin, 2017; Kazanina et al., 2018; Schmitz et al., 2018). In the spoken production of language a distinction can be made between lexical and post-lexical phonological processes and representations (Goldrick and Rapp, 2007; Goldrick, 2014). For example, in order to express a concept, two broadly distinguishable stages of processing are generally assumed to be engaged. First, an abstract lexical-phonological representation (i.e. word-form) that lacks some of the dimensions of phonological information is retrieved from long-term memory (i.e. lexical level). Second, the corresponding phonemic and syllabic representations of phonological structure are subsequently integrated with the lexical-phonological representation to encode a coherent and more fully specified post-lexical phonological representation (i.e. post-lexical level) that can serve as input to further articulatory and motor processing.

More specifically, in naming the picture of a “cap” it is commonly agreed that the image is visually processed until the object is recognised, which then activates the semantic (e.g., features such as <clothing, academic, head covering>) and syntactic representations (e.g., <noun>) associated with the concept depicted by the picture. Next, an abstract word-level representation (or lexical representation) is selected (e.g., <CAP>) that influences the retrieval of a lexical-phonological representation (i.e. phonological retrieval) specifying the identity and serial order of the phonemes (e.g., /k æ p/). In addition, the interaction of semantic, lexical and phonological levels may lead to the partial activation of semantically- (e.g., <HAT>) and phonologically-related (e.g., <TAP>) lexical competitors (aka neighbours). These in turn cause the retrieval of some of the segment-sized representations corresponding to the partially activated semantic and phonological neighbours of the target word (e.g., Rapp and Goldrick, 2000; Goldrick and Rapp, 2002). The selected lexical-phonological representation serves as input to processes that further specify the content of the syllabic (e.g., <CVC>) and featural information of the encoded phonological structure (e.g., /k^h æ p/ indicating that the initial /k/ in ‘‘cap’’ is aspirated due to the context in which is embedded; but see Roelofs, 1997

and Levelt et al., 1999). The ensuing fine-grained phonological representation is subject to additional phonetic processing to generate articulatory representations that drive the articulators (Browman and Goldstein, 1992; Goldrick and Blumstein, 2006; Ziegler et al., 2010, 2012; Buchwald and Miozzo, 2011; Laganaro, 2012; Buchwald, 2014; Maas et al., 2014; Galluzzi et al., 2015).

When repeating words out loud, the acoustic input has to be translated into a phonological code (i.e. phonological recoding) prior to articulation, meaning that an important component of repetition involves the successful recognition of phonemes in addition to the retrieval and mental manipulation of phonological representations. Importantly, performance on word repetition tasks has been shown to dissociate from that on non-word repetition tasks (non-words are made-up words like “clup” and can be thought of as being exposed for the first time to an unfamiliar or foreign word) (Shallice, 1988; Ellis and Young, 1996). Moreover, neuropsychological studies have documented that lexical-semantic and phonological processing abilities can independently be affected by brain damage as indicated by the performance/error patterns seen in patients during speech reproduction (and other language) tasks (Michel and Andreewsky, 1983; McCarthy and Warrington, 1984, 2001; Butterworth and Warrington, 1995; Ingles et al., 1996), which suggests the existence of at least two alternative pathways for repetition (and production more generally). Non-word repetition is usually assumed to take place via a sub-lexical (or non-lexical or phonological) route that deals with processing units smaller than whole words by mapping acoustic input directly onto representations of phonological structure, thereby bypassing the lexical level (because non-words have no associated lexical or semantic entries in long-term memory) (e.g., Patterson and Shewell, 1987; Hillis and Caramazza, 1991; Hanley et al., 1997, 2002). Word repetition, on the other hand, can be accomplished via either the lexical-semantic or sub-lexical routes depending on individual preferences and/or stimulus characteristics (e.g., imageability, frequency, etc.). The output of either of these recoding steps is considered

to be subject to further post-lexical phonological and phonetic processing before speech production can occur (Goldrick and Rapp, 2007).

On the basis of evidence from brain-damaged patients (e.g., Rapcsak et al., 2007), similar functional organisational principles have also been proposed for word and non-word reading with the exception that the reading of non-words would entail an orthographic-to-phonological recoding level (i.e. sub-lexical pathway) instead of an acoustic-to-phonological recoding level (Coltheart et al., 2001; Coltheart, 2005). Crucially, from a connectionist point of view (e.g., Dell et al., 2007; Seidenberg, 2012), the phonological errors observed during word and non-word reading would arise from disruption to the same phonological system involved in speech production and perception (i.e. reflect a modality-independent phonological impairment) rather than from damage to modality-specific mechanisms (Crisp and Lambon Ralph, 2006; Jefferies et al., 2007; Rapcsak et al., 2009; but see Tree and Kay, 2006).

1.6.3 Phonological working memory

Once phonological representations have been retrieved from long-term memory, it is typically necessary to maintain these phonological traces in short-term memory (i.e. a temporary storage of information) while mental operations are performed and further processes are engaged in response to the processing demands of the task at hand (Baddeley, 1996). Such a cognitive device is commonly referred to as phonological working memory. One of the most influential theories of working memory is the three-component system proposed by Baddeley and Hitch in 1974 (but see Cowan, 2008, 2010) which posits the existence of the following three interacting modules: (1) the phonological loop (aka the articulatory loop); (2) the visuo-spatial sketchpad; and (3) the central executive. According to this framework, the central executive represents an attentional control system and is supported by two subsidiary slave systems that specialise in holding phonological and visuo-spatial information (i.e. the phonological loop and the visuo-spatial sketchpad). The model was later refined by adding a fourth component (i.e. the episodic buffer) in an attempt to account for a range of phenomena

that was not satisfactorily captured by the original version (Baddeley, 2000). This episodic buffer is assumed to be a limited capacity store capable of binding information from long-term memory and the subsidiary working memory systems into a unitary multidimensional episodic representation.

Understanding the phonological loop is most relevant for the goals of the current thesis. It has been shown to be particularly important for the acquisition of novel vocabulary during development or when learning a new language (Baddeley et al., 1988; Gathercole and Baddeley, 1989; Service, 1992) and is also likely to be important for recovering phonological processing abilities after stroke. The phonological loop is thought to comprise two sub-systems: (i) a phonological store (or buffer) that temporarily maintains memory traces of verbal material in a phonological code (Schweppe et al., 2011); and (ii) a subvocal rehearsal mechanism that prevents the memory traces of phonological information from decaying with time through covert articulation (Baddeley, 2012). Yet, based on differential impairments seen in brain-damaged patients performing phonologically demanding tasks, it has been proposed that the phonological loop component of working memory may in reality consist of two temporary memory stores (Nickels et al., 1997; Martin et al., 1999; Howard and Nickels, 2005): (1) a phonological input buffer which holds a pre-lexical phonological code and (2) a phonological output buffer which holds a post-lexical phonological code that is required for repeating, reading and writing (to dictation) words and non-words (Caramazza et al., 1986; Shallice et al., 2000). In the context of these two buffers, rehearsal would be achieved via the cycling of information between input and output. Critically, there is evidence for common underlying processes between phonological working memory and phonological processing (Martin and Saffran, 1997; see also Gupta and MacWhinney, 1997 and Martin and Saffran, 2002); a point I will come back to later on in the Introduction.

1.7 Neurobiology of phonological processing

1.7.1 Neurobiology of phonological processing in neurologically-normal individuals

Many functional neuroimaging studies of the neurologically-normal brain have shown that activity in a left-lateralised network of regions increases with increasing demands on phonological processing. These brain regions include: the supramarginal gyrus (e.g., Price et al., 1997; Gelfand and Bookheimer, 2003; Moser et al., 2009; Kircher et al., 2011; Oberhuber et al., 2016), posterior part of the inferior frontal gyrus (e.g., Heim et al., 2003; Burton et al., 2005; Papoutsi et al., 2009; Vaden et al., 2011; Okada et al., 2018), premotor cortex (e.g., Devlin et al., 2003; McDermott et al., 2003; Katzir et al., 2005; Mechelli et al., 2005a; Twomey et al., 2015) and mid-posterior portion of the superior temporal gyrus (Buchsbaum et al., 2001; Graves et al., 2008; McGettigan et al., 2011; Arsenault and Buchsbaum, 2015). For example, in an elegant study, Peramunage et al. (2011) manipulated phonological neighbourhood density (i.e. the number of words in the mental lexicon that deviate from a target word by a single phoneme) during word production to show that activity in the left premotor cortex (PMC), inferior frontal gyrus (IFG), posterior superior temporal gyrus (pSTG) and supramarginal gyrus (SMG) responded to phonological competition; indicating that these regions form part of a network involved in accessing, maintaining and manipulating the phonological properties of words. Similarly, Bohland & Guenther (2006) visually presented neurologically-normal participants with syllable sequences of varying phonological complexity to reveal that a collection of brain regions including PMC, IFG, STG and SMG is important for planning and producing sequences of speech sounds. Subsequently, it has been demonstrated that the distributed patterns of activity in these regions encode phonological information at the phonemic and/or syllabic levels (Markiewicz and Bohland, 2016; see also Peeva, 2010) or even lower-level features such as place/manner of articulation and voicing (Correia et al., 2015).

Consistent with an involvement in phonological processing, this dorsal stream of language regions is also engaged during reading (e.g., Jobard et al., 2003; Graves et al., 2010; Taylor et al., 2013; Danelli et al., 2015; Malins et al., 2016), verbal working memory (e.g., Paulesu et al., 1993; Awh et al., 1996; Smith et al., 1998; Buchsbaum et al., 2005; Perrachione et al., 2017), repetition (e.g., Price et al., 2003; Shuster and Lemieux, 2005; Saur et al., 2008; Yoo et al., 2012; Hope et al., 2014), writing-to-dictation (e.g., DeMarco et al., 2017) and other tasks which heavily rely on the successful retrieval and mental manipulation of phonological representations at lexical and sub-lexical levels. Moreover, the vPMC, pIFG, SMG and pSTG have been found to contribute to the categorical perception of speech sounds (Raizada and Poldrack, 2007; Lee et al., 2012; Chevillet et al., 2013; Du et al., 2014; Bouton et al., 2018).

Indeed, dorsal language regions are posited to subserve the mapping between sensory and motor phonological codes for overt as well as covert speech processing (e.g., Hickok and Poeppel, 2007; Saur et al., 2008); see Figure 1.1. For instance, in an illuminating study of the mental imagery of speech, Tian et al. (2016) found that speech perception and production are strongly linked at the neural level (see also Skipper et al., 2017) and that covert and overt speech recruit overlapping brain networks including the opercular part of the left inferior frontal gyrus and the anterior part of the left supramarginal gyrus which strongly suggests the availability of shared phonological representations. Furthermore, naturalistic phonological manipulations between speech perception and speech production have revealed that segmental processing (i.e. computations that operate at the phonemic level) during inner speech activates left-lateralised dorsal speech regions such as pIFG and SMG (Peschke et al., 2012). In line with these results, it has been suggested that covert articulation or inner speech (for recent reviews, see Perrone-Bertolotti et al., 2014 and Alderson-Day and Fernyhough, 2015) carries detailed phonological information and may ultimately reflect a special type of overt speech (Corley et al., 2011; Schweppe et al, 2011; Niziolek et al., 2013; Whitford et al., 2017; but see Oppenheim and Dell, 2010).

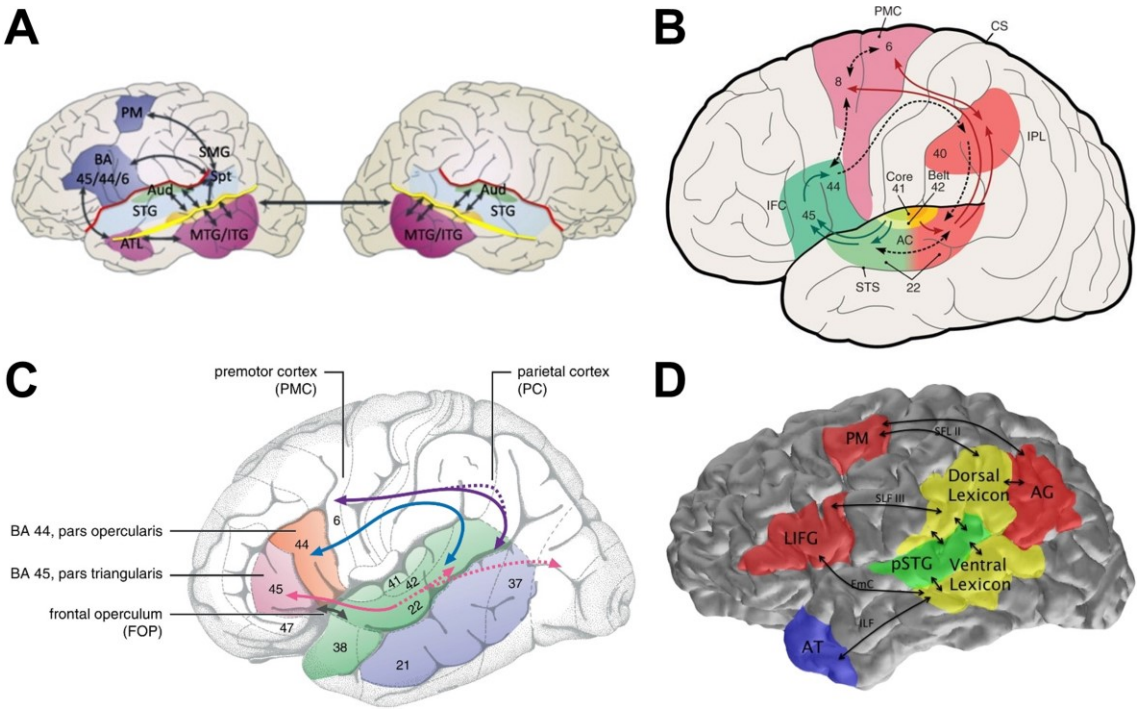

Figure 1.1: Dual-stream models of the language network in the human brain. The figure provides a schematic depiction of four alternative implementations of the dual-stream theory of speech processing taken from (A) Hickok (2009), (B) Rauschecker & Scott (2009), (C) Friederici & Gierhan (2013) and (D) Gow (2012). In the context of the dual-stream framework, it is proposed that motor-phonological aspects of speech processing are mediated by a dorsal pathway, whereas lexical-semantic aspects of speech processing are supported by a ventral pathway as indicated by ample experimental evidence from both neurologically-normal and brain-damaged individuals (e.g., Saur et al., 2008; Kümmerer et al., 2013; Mirman et al., 2015; Fridriksson et al., 2016). Syntactic aspects of speech processing would, on the other hand, load more evenly on both these pathways (e.g., Rolheiser et al., 2011; but see Wilson et al., 2011). In general terms, the dorsal route is thought to involve a set of regions including the mid-posterior part of the superior temporal gyrus, supramarginal gyrus, pars opercularis of the inferior frontal gyrus and premotor cortex; all connected via branches of the superior longitudinal fasciculus and/or arcuate fasciculus. In contrast, less agreement exists with regard to the components of the ventral route which may incorporate the full set or a subset of the following: mid-anterior part of the superior temporal gyrus, middle temporal gyrus, inferior temporal gyrus and pars triangularis of the inferior frontal gyrus; all connected via the middle longitudinal fasciculus, inferior longitudinal fasciculus, uncinated fasciculus and extreme capsule. Another point of disagreement relates to the (i) degree of lateralisation and (ii) specific computational properties of the two streams. For more details (including key to abbreviations), the reader is referred to the original publications. Numbers correspond to Brodmann areas. Adapted with permission from Elsevier and Springer Nature.

Causal evidence favouring the view that dorsally located language areas (e.g., left pIFG, PMC, SMG and pSTG/STS) contribute to phonological processing has been provided by methods that interfere with the neural activity in selected brain regions (e.g., Sato et al., 2009; Acheson et al., 2011; Murakami et al., 2015; Hartwigsen et al., 2016).

This has led a substantial number of researchers to implicate these brain regions in a range of phonological processes including: (i) the categorical perception of phonemes (Meister et al., 2007; Krieger-Redwood et al., 2013); (ii) phoneme segmentation (Sato et al., 2009); (iii) phonological working memory (Herwig et al., 2003; Nixon et al., 2004; Romero et al., 2006); (iv) phonological recoding (Nakamura et al., 2006); and (iv) phonological encoding/retrieval (Acheson et al., 2011; Savill et al., 2018). Transcranial magnetic stimulation (TMS) studies have therefore provided converging evidence which demonstrates that regions such as the left anterior supramarginal gyrus and the opercular part of the left inferior frontal gyrus are important nodes in the dorsally distributed phonological network (e.g., Gough et al., 2005; Hartwigsen et al., 2010a, b; Sliwinska et al., 2015).

Regarding the temporal dynamics of phonological processing in the dorsal language pathway, the evidence accumulated thus far points to a highly interactive perisylvian system that operates at multiple spatial and temporal scales (e.g., Munding et al., 2016). For instance, Liebenthal et al. (2013) simultaneously acquired event-related potentials and functional magnetic resonance imaging (fMRI) data and found that the discrimination of duplex syllables (composed of chirp and base portions) presented dichotically at varying interaural asynchronies recruit a set of regions comprising the posterior superior temporal gyrus bilaterally, left supramarginal gyrus, left ventral postcentral (vPostCG) and precentral gyri (vPreCG). More importantly, during syllable identification (compared to chirp identification) activity arose in bilateral pSTG at ~80 ms (peaking at ~100 ms), closely followed (~90 ms) by the engagement of the left SMG (peaking at ~140 ms), left vPostCG and left vPreCG (both exhibiting similar temporal profiles to left SMG); with activity reverberating in a highly overlapping network after 300 ms from stimulus onset. Crucially, the elicited activity was significantly left-lateralised in SMG, vPostCG and vPreCG in the early time window, and in pSTG only in the late time window (> 300 ms), which may be indicative of interactive processing with efferent feedback from sensorimotor to auditory cortex. In another insightful study, Herman et al.

(2013) recorded magnetoencephalographic responses while neurologically-normal individuals listened to and reproduced (after a brief delay) two- or four-syllable utterances to show that the phonological loop component of verbal working memory is underpinned by regions within the dorsal speech stream. By correlating neural oscillations and performance, they found that phonological encoding, maintenance and response preparation (prior to overt articulation) recruit a primarily left-lateralised network where information cycles between input and output buffers in temporal/parietal (i.e. pSTG/SMG) and frontal/premotor (i.e. pIFG/vPMC) areas while a coherent phonological representation is formed as part of a reverberant process.

1.7.2 Neurobiology of phonological processing in clinical populations

In agreement with the findings from fMRI and TMS investigations (summarised above), studies of clinical populations have contributed additional evidence to support the view that phonological processing is instantiated in dorsal language regions (e.g., Kümmerer et al., 2013; Mirman et al., 2015; Fridriksson et al., 2016). For example, Schwartz et al. (2012) found that the rate of phonological errors produced during picture naming is associated with stroke damage to various components of the left-lateralised dorsal language stream including but not limited to pIFG, PMC, PreCG, PostCG, SMG and STG. Likewise, phonological performance (as indexed by phonological manipulation, reading and writing-to-dictation tasks of both words and non-words) in patients with primary progressive aphasia (i.e. a neurodegenerative syndrome) has been shown to depend upon the structural integrity of the grey matter in frontal and parietotemporal regions of the left hemisphere including the pIFG, PMC, PreCG, PostCG, SMG and STG (Henry et al., 2016).

Intraoperative mapping in large series of patients undergoing awake surgery for resection of low-grade gliomas has confirmed that phonological/phonetic errors elicited during picture naming arise from stimulation of dorsal speech regions (Tate et al., 2014; see also Roux et al., 2012). Furthermore, intraoperative electrical stimulation over the left pIFG and SMG in tumour patients has been shown to disrupt phonological working

memory performance (Papagno et al., 2017). Linking phonological input and output processing, Cogan et al. (2014) obtained electrocorticographic (ECoG) recordings from patients with pharmacologically resistant epilepsy and demonstrated that the mapping between sensory and motor phonological representations of speech sounds (i.e. sensory-motor integration) for both perception and production occurs bilaterally in inferior frontal, premotor, motor, somatosensory, inferior parietal and superior temporal cortices. Moreover, the relationship between the overt and covert reproduction of words in the visual and auditory modalities has also been characterised using ECoG signals from patients with intractable epilepsy (Pei et al., 2011). In brief, the results revealed that the spatial topography and temporal course of the neural responses induced by the overt and covert reproduction of aurally and visually presented words largely overlap in inferior frontal, premotor, motor, somatosensory, inferior parietal and superior temporal areas, with overt word reproduction associated with stronger and more distributed cortical activation relative to covert word reproduction (see also Flinker et al., 2015 and Brumberg et al., 2016).

In addition to the documented role of grey matter regions, patient studies have been essential for establishing that dorsal (but not ventral) white matter tracts such as the arcuate fasciculus (AF) and superior longitudinal fasciculus (SLF) support phonological processing (e.g., Friederici and Gierhan, 2013; Parker Jones et al., 2014; Han et al., 2016). For instance, by analysing neuroimaging and behavioural data from a large sample of acute left-hemisphere stroke patients, Kümmerer et al. (2013) showed that poorer performance on tasks that load heavily on phonological processing (such as language repetition) is associated with greater damage to dorsal fibre tracts (i.e. AF/SLF), whereas poorer performance on tasks that load heavily on semantic processing (such as language comprehension) is associated with greater damage to ventral fibre tracts (e.g., the extreme capsule or EmC). Similarly, Rolheiser et al. (2011) studied a sample of 24 chronic left-hemisphere stroke patients who were subject to detailed behavioural testing and neuroimaging examination to find that input/output phonology

correlated with fractional anisotropy (i.e. a proxy measure of white matter integrity) in the dorsal pathway (i.e. left AF), while input/output semantics correlated with fractional anisotropy in the ventral pathway (i.e. left EmC). Moreover, direct electrostimulation of the AF/SLF in patients undergoing tumour surgery induces phonological/phonetic errors (Maldonado et al., 2011) and phonological working memory errors (Papagno et al., 2017) but not semantic errors, consistent with proposals that these white matter tracts form part of the dorsal phonological system. Complementary results have been reported by studies of patients with primary progressive aphasia (e.g., Galantucci et al., 2011). Evidence pointing to the existence of feedforward and feedback projections between left frontal and parietotemporal language areas comes from intraoperative recordings of cortico-cortical evoked potentials in patients with intractable epilepsy which have indicated that the flow of information between pIFG/PMC and SMG/pSTG sites is bidirectional (probably through the AF/SLF) (Matsumoto et al., 2004). These findings, in conjunction with recent methodological advances (e.g., Jones et al., 2013), have led to a renewed interest in tracing the connectional anatomy of language and to a fuller appreciation of the critical role that white matter fibre bundles play in transferring information between functionally related but spatially distant nodes of the same brain network (e.g., Martino et al., 2013; Fernández-Miranda et al., 2015; Wang et al., 2016; Yagmurlu et al., 2016).

1.7.3 Phonological working memory as an emergent property of the language system

Traditionally, phonological working memory and language processing have been thought to constitute two independent systems that can be studied separately (e.g., Baldo and Dronkers, 2006; Baddeley and Hitch, 2018). However, more contemporary proposals conceive phonological working memory as a more fundamental property of the speech perception and production systems (e.g., Jacquemot and Scott, 2006; Buchsbaum and D'Esposito, 2008; Acheson and MacDonald, 2009a; Majerus, 2013) based on four empirical observations: (i) the short-term maintenance of linguistic information relies on underlying language processes (e.g., Martin et al., 1994; Martin and

Saffran, 1997, 1999; Knott et al., 2000); (ii) serial position errors during phonological working memory tasks are better described as phoneme rather than whole item ordering errors (e.g., Nimmo and Roodenrys, 2004; Page et al., 2007; Acheson and MacDonald, 2009b; see also Majerus et al., 2004); (iii) brain regions activated by the perception and production of speech are also implicated in phonological working memory (e.g., Acheson et al., 2011; Ravizza et al., 2011; Perrachione et al., 2017); (iv) speech processing and verbal working memory performance are disrupted following TMS-induced "virtual" lesions (e.g., Acheson et al., 2011) or stroke-induced "real" lesions (e.g., Leff et al., 2009; Koenigs et al., 2011; Baldo et al., 2012; Hickok et al., 2014) to the same brain regions (see also Meyer et al., 2015).

In particular, Martin & Saffran (1997) have provided evidence that the processes that govern the activation and maintenance of phonological representations for phonological processing and phonological working memory may be common to both by showing that, in brain-damaged patients, behavioural performance on verbal working memory tasks co-vary with that on more general language tasks. Moreover, an important role for sub-lexical phonological encoding in maintaining serial order information and a close link between the mechanisms supporting phonological processing and phonological working memory have been proposed by Acheson & MacDonald (2009b). They manipulated phonological similarity of lists of non-words by prompting neurologically-normal individuals to perform a variety of verbal working memory tasks involving tongue twisters matched for overall phonological overlap and found that the serial order errors produced by the participants due to phonological similarity followed phonological constraints on the production system (such as syllable-position constraints). In line with these findings, Acheson et al. (2011) concluded that the capacity to hold sequences of speech sounds over short periods of time is contingent upon representations within the phonological network after demonstrating that TMS over a region within the pSTG associated with phonological encoding increased error rates on paced reading and delayed serial recall of non-words but not on picture naming.

Similarly, patient studies have revealed that repetition and phonological working memory abilities as measured by non-word repetition and digit span tasks (both of which place high demands on phonological processing) overlap at the neural level, with impairments on these tasks resulting from damage to the left mid-posterior superior temporal/inferior parietal cortex (e.g., Baldo et al., 2012). The evidence therefore points to the existence of a shared functional architecture between phonological processing and phonological working memory.

In this context, phonological working memory is thought to (a) emerge from the sustained activation of abstract sub-lexical motor and sensory speech codes via dynamic motor-to-sensory (or feedforward) and sensory-to-motor (or feedback) mappings during covert speech as indicated by neuromagnetic responses recorded from neurologically-normal subjects during the covert rehearsal of aurally presented multisyllabic non-words (Ylinen et al., 2015; see also Cogan et al., 2017) and (b) be instantiated in dorsal phonological regions within the inferior frontal/premotor and inferior parietal/superior temporal cortices involved in both the production and perception of speech as revealed by prior functional imaging (e.g., Paulesu et al., 1993), transcranial magnetic stimulation (e.g., Romero et al., 2006) and lesion studies (e.g., Baldo and Dronkers, 2006), probably in tight interaction with other language and cognitive systems (e.g., Ruchkin et al., 2003; Awh et al., 2006; Linden, 2007; D'Esposito and Postle, 2015; Eriksson et al., 2015).

1.7.4 Neurocomputational models of speech processing

While functional neuroimaging studies of neurologically-normal individuals in combination with brain mapping studies of neurological patients are essential for identifying the brain regions that represent key features of the speech signal and its neural dynamics, neurobiologically plausible computational models are needed to integrate current scientific knowledge into large-scale functional architectures that provide a mechanistic description of how the neural computations underlying speech are instantiated and interact across the brain (Teufel and Fletcher, 2016; Kriegeskorte and Douglas, 2018). In addition, computationally-implemented theories of speech production

are particularly useful for generating very specific hypotheses that can be tested in future experiments so as to gather new evidence and refine existing models. Two influential examples of such a methodological approach are the Directions Into Velocities of Articulators (DIVA) model (Golfinopoulos et al., 2010; Tourville and Guenther, 2011; Guenther and Vladusich, 2012) and the Hierarchical State Feedback Control (HSFC) model (Hickok, 2012, 2014; see also Houde and Nagarajan, 2011).

One important commonality between these models is that they both assume the existence of independent but linked motor and sensory components of speech sound (phonological) representations in frontal and parietotemporal cortices that are activated in parallel during covert and overt speech production as indicated by compelling experimental evidence (e.g., Jacquemot et al., 2007; Golfinopoulos et al., 2011; Lametti et al., 2012; Evans and Davis, 2015; Ding et al., 2016). The DIVA model consists, in general terms, of feedforward and feedback control systems that underlie speech production, with the feedback control system incorporating neural assemblies in the left pIFG/vPMC, SMG and pSTG that represent motor (aka “speech sound map”), somatosensory (aka “somatosensory target map”) and auditory (aka “auditory target map”) information related to discrete segments of speech, typically encoded at the syllable level. Similarly, the HSFC model posits that a feedback detection/correction mechanism links abstract motor, somatosensory and auditory representations of speech sounds in pIFG/vPMC, aSMG and STG/STS. In addition, both neurocomputational architectures embody in their feedback “phonological” loops forward and inverse models that predict the sensory consequences of articulatory gestures and transform prediction errors into correction signals (for more details, see Wolpert and Ghahramani, 2000 and Pickering and Clark, 2014).

The HSFC model is not identical, however, to the DIVA model in that the HSFC model (i) includes an internal feedback control mechanism that can detect and correct errors prior to any overt speech production, (ii) assumes a hierarchical organisation of the auditory and somatosensory feedback mechanisms where the auditory feedback

operates at the syllable level (Assaneo and Poeppel, 2018) and the somatosensory feedback operates at the phoneme level (but see Sato et al., 2014), and (iii) entails a modified computational architecture for the auditory-motor and somatosensory-motor interfaces (Guenther and Hickok, 2016). Conversely, the DIVA model has been extended to account for the planning and production of simple speech sequences (Bohland et al., 2010).

With respect to computational architectures of both speech perception and production, Ueno et al. (2011) developed a neurobiologically grounded model of the dual dorsal-ventral language pathways and trained it to name, repeat and comprehend hundreds of multi-syllabic Japanese words. Neuroanatomically, the neural network comprised (1) a dorsal pathway with each layer mapping onto specific cortical regions such as primary auditory/pSTG, inferior supramarginal gyrus and insular-motor cortices, all connected by the arcuate fasciculus and (2) a ventral pathway with each layer representing cortical structures such as primary auditory/pSTG, mid-STG, anterior-STG, opercularis-triangularis and insular-motor cortices, all connected by the middle longitudinal fasciculus and the extreme capsule. Critically, a dual-pathway architecture compared to a ventral-only architecture proved to be more efficient at learning the differential patterns of performance on language production and comprehension tasks (but see Roelofs, 2014). The results also showed that, during the training process, the model adopted a partial division of labour such that damage to the dorsal stream had a greater effect on single-word repetition whereas damage to the ventral stream had a larger impact on single-word comprehension and naming. Interestingly, by lesioning specific components of the network, the authors were able to simulate various forms of vascular and progressive aphasias, with recovery of language function arising (at least partially) as a function of increased reliance on the intact pathway.

On the other hand, an influential connectionist model of language production is the dual-route interactive two-step model which has been widely used to replicate variations in the performance of neurologically-normal subjects and brain-damaged

patients during picture naming, word repetition and non-word repetition (e.g., Schwartz and Dell, 2016; Tochadse et al., 2018). The computational architecture of the model comprises three interconnected layers containing nodes that code semantic, lexical, input and output phonological features of words (Nozari et al., 2010). More specifically, the free parameters (those that are fitted to actual data) are the bidirectional connection weights between the nodes in each layer: (i) the ***s*** parameter between semantic and lexical units, (ii) the ***p*** parameter between lexical and output phonological units, and (iii) the ***nl*** parameter between input and output phonological units. In a recent study, Dell et al. (2013) devised a new lesion-deficit mapping approach by relating the model parameters to the brain. In particular, the error patterns of a large sample of 103 stroke patients naming pictures and repeating non-words were modelled to obtain the *s*, *p* and *nl* connection weights and then these 3 parameters were mapped onto the brain. The neural correlates of the *p*-weight involved parts of the dorsal stream including the SMG, postcentral gyrus, precentral gyrus, and insula. In turn, the lesion-parameter mapping of the *nl*-weight identified substantial portions of the STG, the posterior third of the planum temporale and cortex at the juncture of the parietal and temporal lobes as well as the SMG and postcentral gyrus. The high degree of overlap between the brain regions associated with the *p*- and *nl*-weights led the authors to revise the model and hypothesise that the function served by these parameters may correspond to a system that links sensory and motor phonological representations of speech.

1.8 Lesion-deficit mapping studies

The emergence of neurocomputational models of phonological processing offers model-based predictions for how stroke damage to the phonological system will affect phonological processing abilities. Before describing how I tested specific hypotheses related to the supramarginal gyrus (SMG) and the pars opercularis (pOp) of the inferior frontal gyrus, I discuss other approaches for mapping lesion sites to deficits. Both these alternative approaches entail using phonological processing abilities as an independent (known) variable and then searching for brain regions where damage is greater in those

who are more impaired (i.e. lesion site is the dependent variable). The lesion search can either be performed at each brain voxel independently (univariate analyses) or by considering combinations of regions (multivariate analyses). I discuss each in turn, highlighting their respective limitations and explaining why mapping from deficit-to-lesion does not necessarily imply that lesions can be mapped to deficits. This is connected with the distinction between "inference" and "prediction" (Shmueli, 2010) and is very important in relation to predicting outcome after stroke when we know the lesion site (the independent variable) but not how it will affect phonological processing abilities and their recovery over time (the dependent variable).

1.8.1 Interpreting univariate voxel-based lesion-deficit mappings

Voxel-based lesion-deficit mapping has become a widely accepted methodology for identifying which brain regions are responsible for carrying out specific cognitive functions (Rorden and Karnath, 2004). For example, voxel-based morphometry (VBM) (Ashburner and Friston, 2000; Mechelli et al., 2005b) has been applied to detect structural brain abnormalities in a broad range of neurological and psychiatric disorders (e.g., Rosen et al., 2002; Gorno-Tempini et al., 2004; Honea et al., 2005; Radua and Mataix-Cols, 2009). VBM (implemented in Statistical Parametric Mapping or SPM) can also be used to conduct voxel-based multiple regression lesion analyses thanks to the flexibility of the general liner model (Tyler et al., 2005; Price et al., 2010b). A similar technique, voxel-based lesion-symptom mapping (VLSM), has also been used to map lesion sites to functional impairments (Bates et al., 2003; Rorden et al., 2007, 2009). Crucially, in all variants of voxel-based lesion-deficit mapping, the association between structural abnormality (the lesion) and functional impairment (the deficit) is measured separately in thousands of very small 3-dimensional volumetric brain units called voxels; and significant lesion-deficit relationships are reported after correcting for the number of statistical tests conducted (for a comparison of VBM and VLSM, see Geva et al., 2012).

Univariate voxel-based lesion-deficit analyses have made significant contributions to uncovering the neural bases of multiple cognitive functions (e.g., Saygin,

2007; Moro et al., 2008; Verdon et al., 2010; Ionta et al., 2011; Barbey et al., 2012; Gläscher et al., 2012; Manuel et al., 2013; Azuar et al., 2014; Melloni et al., 2016; Halai et al., 2017; Martin et al., 2017). However, there are a number of challenges that need to be considered: first, statistically significant group-level lesion-deficit mappings do not necessarily indicate consistent lesion effects across individual patients. Indeed, my colleagues and I recently studied a very large sample of more than 300 right-handed left-hemisphere stroke patients who collectively acquired a wide range of lesion sizes and locations to identify the brain regions where damage was associated with word finding difficulties (Gajardo-Vidal, Lorca-Puls et al., 2018). We found two spatially-distinct clusters in frontal and temporal regions where structural abnormality was highly significantly associated with the language function being investigated. The first cluster extended from the posterior middle temporal gyrus into the arcuate fasciculus, temporal stem and anterior superior temporal gyrus. The second cluster was centred on the left ventral premotor cortex including the surrounding white matter. Critically, however, post-hoc analyses showed that damage to either the frontal or temporal regions resulted in word finding difficulties in less than 50% of the affected patients. These results therefore clearly demonstrate that inferential methods that rely on the statistical significance of group-level effects are not well suited for generating individualised outcome predictions, unless further post-hoc tests indicate consistency in the inter-subject lesion-deficit mapping prior to validation (see also Price et al., 2017 and Halai et al., 2018).

A second challenge with interpreting the results from voxel-based lesion-deficit analyses is that statistically significant group-level effects in the context of inter-subject variability do not replicate across studies. For example, my co-workers and I characterised the impact of sample size on the reproducibility of univariate voxel-based lesion-deficit mappings (Lorca-Puls et al., 2018). Using a total “population” of 360 right-handed left-hemisphere stroke survivors, we investigated how the strength and statistical significance of the same lesion-deficit mapping (i.e. within a pre-defined region of interest) varied with sample size. When the analysis included 30 patients, we found that

the mean effect size (expressed in R^2 terms) across hundreds of different samples (all N = 30) that yielded significant results ranged from 0.16 to 0.79 (M = 0.26); but when the analysis included 180 patients, the significant results were associated with much smaller effect sizes with R^2 ranging from 0.02 to 0.38 (M = 0.12). Critically, using a more stringent statistical threshold only aggravated the problem because when we changed the alpha level from $p < 0.05$ to $p < 0.001$ the mean effect size of the significant resamples when N = 30 became 0.45 (range = 0.38-0.79) compared to 0.12 (range = 0.06-0.38) when N = 180.

As expected, statistical power was also extremely sensitive to sample size as reflected by the fact that significant results were observed for 36.9% of samples that included 30 subjects relative to 99.6% for samples that included 180 subjects (at an uncorrected statistical threshold of $p < 0.05$ in the region identified from the analysis on all 360 patients). This set of findings illustrates how inter-patient variability in the effect of the same lesion sites poses non-trivial challenges for translating univariate lesion-deficit mappings into accurate outcome predictions. It also highlights the importance of studying very large samples of patients in order to be able to understand and model the main sources of inter-patient variability. For example, one source of inter-subject variability arises when damage to more than one brain region (e.g., regions A or B) can independently cause the same functional impairment (e.g., Gajardo-Vidal, Lorca-Puls et al., 2018). In this case, a group of patients with the same functional impairment may either have damage to: (i) A not B; (ii) B not A; or (iii) A & B. If there are plenty of patients with damage to A & B, then both regions may be detected. But if there are a few patients with damage to A & B and an equal number of patients with damage to A or B, no significant effects would be observed. Conversely, if a function can be sustained by region A or B, then functional impairments may only become apparent when there is damage to a combination of brain regions (e.g., regions A & B) (e.g., Seghier et al., 2014; Pustina et al., 2018). This logic helps to explain why the results of voxel-based lesion analyses vary with the patient sample tested.

A third challenge with interpreting the results from voxel-based lesion-deficit analyses is that the location and extent of stroke lesions do not generally follow the size and shape of the functional units of the brain, but are instead determined by vascular factors (Kimberg, 2007). For example, Mah et al. (2014) modelled inter-voxel dependencies of brain damage in more than 500 acute vascular injuries to reveal that univariate voxel-based lesion-deficit mapping techniques are incapable of dissociating critical voxels from non-critical voxels that have nonetheless been co-incidentally damaged by stroke, resulting in the displacement of lesion-deficit mappings towards areas of greater susceptibility to vascular events. Similar findings have also been reported by Inoue et al. (2014) and Sperber & Karnath (2017). This systematic and pervasive hidden spatial bias within the multivariate patterns of damage distorts the results of univariate voxel-based lesion-deficit maps (towards vascular boundaries) because if, for instance, region A is irrelevant for function X but always damaged in conjunction with region B which is critical for function X, then the critical locus of damage will be mislocalised towards area A.

Together, the points discussed above highlight the importance of studying (i) how the effect of damage to one region depends on that in another; and (ii) how vascular damage after stroke influences the mapping between lesion site and deficit.

1.8.2 Interpreting multivariate machine-learning-based lesion-deficit mappings

Some of the challenges outlined above are starting to be addressed with new methods that (unlike univariate approaches) capture the multivariate patterns in which brain damage affects behaviour (Xu et al., 2018). This novel collection of techniques involves using advanced machine learning algorithms to explicitly model (i) how the effect of damage in one voxel depends on that in another and (ii) the parasitic voxel-to-voxel associations that arise as a result of the non-random distribution of vascular damage across the brain. For instance, Zhang et al. (2014) introduced a non-linear support vector regression-based (SVR) multivariate lesion-deficit mapping approach and showed that, in comparison with a standard univariate approach, it (a) yielded

significantly higher performance accuracy (higher sensitivity and specificity) in detecting ground-truth lesion-deficit relationships generated from synthetic data and (b) was more sensitive to empirical lesion-deficit relationships present in real patient data. A five-fold cross-validation scheme iterated 40 times revealed, however, that the same SVR method only achieved a relatively modest prediction accuracy (i.e. the ability to predict functional outcomes in new patients): R^2 = ~0.10. Likewise, Pustina et al. (2018) demonstrated that a multivariate lesion-deficit mapping method based on sparse canonical correlations produced more accurate lesion-deficit mappings in terms of smaller localisation errors (although far from being perfect) than state of the art univariate implementations in most of the simulated situations including different: (a) sample sizes, (b) corrections for multiple comparisons and (c) multi-area combinations. Moreover, when applied to real patient data, it exhibited higher sensitivity than its univariate counterpart. Nonetheless, the cross-validated (i.e. 4-fold cross-validation scheme) correlation between predicted and observed behavioural scores was still lower than 0.60 (i.e. $R^2 < 0.35$).

Other efforts have placed a greater emphasis on exploiting the promising potential of machine learning algorithms to build data-driven predictive models that can successfully learn the most important multivariate structure-function-recovery rules underlying accurate outcome predictions (e.g., Hope et al., 2015; Rehme et al., 2015a, b; Siegel et al., 2016). For example, Smith et al. (2013) trained and tested a non-linear support vector machine classifier on lesion data from a large sample of 140 right-hemisphere stroke patients to predict the presence or absence of spatial neglect. By adopting a leave-one-out cross-validation scheme, the authors found that, when the classifications examined the contribution of multiple voxel across the whole brain simultaneously, the predictive power of patches of voxels (i.e. multivariate classification) outperformed the best-performing single voxel (i.e. univariate classification). In addition, when the lesion analysis considered the degree of damage to 45 right hemisphere regions, a significantly higher average prediction accuracy was observed for two-region combinations than single regions (lesion-size-adjusted percentages: ~62% versus

~53%); and for three-region combinations than two-region combinations (lesion-size-adjusted percentages: ~67% versus ~62%). Similarly, Hope et al. (2013) trained and tested a Gaussian process regression model on lesion and non-lesion data from a large sample of 270 stroke patients to predict speech production scores in new cases at the individual subject level. The best predictor configuration included time post-stroke, lesion size and lesion load (i.e. the proportion of damaged voxels among all voxels within discrete anatomical structures) in 35 grey/white matter regions, which were selected using a fully automated procedure. Nevertheless, the predicted scores (obtained from a leave-one-out cross-validation scheme) only accounted for 59% of the variance in the observed speech production scores.

The ever-increasing popularity of machine learning algorithms does not come free of its own problems and controversies (e.g., Arbabshirani et al., 2017; Varoquaux et al., 2017; Carlson et al., 2018; Janssen et al., 2018; Mateos-Pérez et al., 2018). In what follows I give seven examples. First, because the number of variables (or feature space) in a typical neuroimaging study is much greater than the number of observations (aka "curse of dimensionality"), machine learning techniques usually incorporate some sort of dimensionality reduction step embedded somewhere in the learning process (Lemm et al., 2011; Klöppel et al., 2012), which could compromise the spatial specificity and interpretability of the results. Second, the choice of input lesion variables is user-dependent (as in univariate methods). However, the shape and size of the functional units of the brain continue to be a matter of debate (Eickhoff et al., 2018a, b; Genon et al., 2018) because cortical areas can be defined on the basis of their structure (e.g., Amunts et al., 1999; Caspers et al., 2006), function (e.g., Sereno et al., 1995; Formisano et al., 2003) and/or connectivity (e.g., Ruschel et al., 2014; Gordon et al., 2016), with distinct brain mapping approaches arriving at different solutions. Consider, for instance, two recent multi-modal parcellations of the human cerebral cortex: the Human Connectome Project atlas (Glasser et al., 2016) and the Brainnetome atlas (Fan et al., 2016). There are 30 more cortical regions per hemisphere in the latter (N = 210) than the

former (*N* = 180). Critically, even small changes in how brain damage is encoded (e.g., at the level of single voxels versus atlas-based regions) lead to noticeable differences in prediction performance (Rondina et al., 2016; see also Abraham et al., 2017). Third, the tuning of the model hyper-parameters (including the regularisation term) is not a trivial issue and does not have a one-size-fits-all solution (Hastie et al., 2004; Lemm et al., 2011; Varoquaux et al., 2017).

A fourth challenge for machine learning users is the lack of consensus as to which type of algorithm to choose when tackling classification (for categorical outcomes) or regression (for continuous outcomes) problems (e.g., Cui and Gong, 2018; Hope et al., 2018). A fifth challenge is that multivariate methods are computationally more expensive than their univariate counterparts, especially if the lesion information is encoded at the voxel level (DeMarco and Turkeltaub, 2018); and could be more complicated to implement in practical terms given the extent of technical knowledge involved. Sixth, due to the high-dimensionality and nonlinearity that arise when attempting to capture multivariate structure-function-recovery associations, the output of machine learning algorithms can be complex and may obscure precise neurobiological interpretations, thereby limiting the degree of scientific insight afforded by these methods (Haufe et al., 2014; Coveney et al., 2016; Huys et al., 2016; Bzdok and Yeo, 2017; Stephan et al., 2017). Seventh, although cross-validation (i.e. splitting the data into training and testing sets) restricts overfitting, the estimated predictive performance of the model is still subject to the vagaries of sample size (Braga-Neto and Dougherty, 2004; Isaksson et al., 2008; Popovici et al., 2010; Cui and Gong, 2018; Varoquaux, 2018). In other words, the use of cross-validation does not preclude the need for proper validation as indicated by studies that reported a substantial drop in effect size estimates after testing the generalizability of the cross-validation results with a full split-half analysis (Price et al., 2013; Pustina et al., 2017).

In summary, while the capacity of multivariate lesion-deficit mapping techniques to model the spatial bias in vascular lesions and the distributed nature of human cognitive

functions is certainly better than that of univariate lesion-deficit mapping techniques, that does not imply that they are perfect. Indeed, Pustina et al. (2018) showed that multivariate methods reduce, but do not completely correct, the displacement of critical areas relative to univariate methods. More research on this topic is thus warranted.

1.9 Thesis overview

In light of the evidence reviewed above, my PhD experiments elaborate and validate a region-of-interest lesion-deficit mapping approach informed by prior TMS and fMRI studies of neurologically-normal individuals. More specifically, I investigate whether regions derived from previous transcranial magnetic stimulation (TMS) and functional magnetic resonance imaging (fMRI) studies of phonological processing in neurologically-normal subjects can be used to guide the search for lesion sites that accurately predict long-term outcome of phonological processing abilities in stroke patients. In other words, the focus of the methodological procedure lies in its ability to identify consistent lesion effects across individual patients rather than the statistical significance of lesion-deficit mappings averaged over multiple patients.

Experiment I sought to determine how well the presence or absence of damage to TMS and fMRI regions of interest predicts the incidence of phonological processing impairments after stroke. Experiment II, on the other hand, tested whether the ability to accurately predict phonological outcomes could be improved by adapting the borders of the regions of interest to include the surrounding grey and white matter where the presence/absence of damage is consistently associated with the presence/absence of phonological processing impairments in stroke patients. Having redefined the borders of the critical lesion sites, Experiment III sought to compare the predictive power of the refined regions of interest with that of a set of regions derived from univariate lesion-deficit mapping techniques. Finally, the goal of Experiment IV was to characterise in detail the functional role of the refined regions of interest by examining patients' performance across a wide range of language tasks that systematically vary the demands on auditory/visual perception, phonology, semantics and speech articulation.

CHAPTER 2:

General Materials and Methods

The purpose of this chapter is to introduce my choice of methodology. First, I explain the main characteristics of the database from which the patient data reported here were retrieved. Second, I describe the standard behavioural assessment used for all PLORAS patients. Third, I delineate additional tasks utilised to link the PLORAS assessment to prior fMRI and TMS studies. Fourth, I provide a brief overview of the fundamental physical principles required to understand magnetic resonance imaging (MRI) from a conceptual viewpoint. Further information can be found in the many exceptional textbooks that have been written about MRI such as those by Poldrack et al. (2011) and Huettel et al. (2014). Fifth, I detail how the structural brain images were acquired and processed giving special attention to the steps involved in the automated definition and delineation of stroke lesions. While the current thesis does not report any data collected using functional MRI (fMRI) or transcranial magnetic stimulation (TMS), in order to fulfil its primary goal it does evaluate the predictive power of regions of interest informed by previous fMRI and TMS studies of neurologically-normal individuals. Hence, the sixth and seventh sections summarise the neurophysiological basis of the fMRI blood-oxygenation-level-dependent (BOLD) response and the biophysical foundations underlying TMS, respectively. The chapter concludes by explaining how the prediction accuracy of the regions of interest will be empirically ascertained.

2.1 The PLORAS database

All the behavioural and structural neuroimaging data reported throughout the current thesis were retrieved from the Predicting Language Outcome and Recovery After Stroke (PLORAS) database (Seghier et al., 2016). No fMRI data from either stroke patients or neurologically-normal controls acquired as part of the PLORAS project are analysed or indeed discussed in subsequent sections. The reader is referred to the experimental chapters where the study-specific patient selection criteria are specified.

The PLORAS database includes brain images, demographic information and the results of behavioural assessments that have been collected from hundreds of stroke patients at the Wellcome Centre for Human Neuroimaging since 2003 with the goal of

understanding, characterising and modelling the most important sources of inter-patient variability in lesion-outcome associations (Price et al., 2010a). Patients can take part in the project independently of the extent and location of the brain damage they have acquired, the presence or absence of aphasia and the amount of time elapsed since stroke onset. Inclusion criteria to the PLORAS database include: (i) a demonstrable previous medical history of stroke; (ii) no record of concomitant neurological or psychiatric illness (e.g., dementia or depression); and (iii) being able to provide written informed consent.

The structural brain scans are run through an automated lesion identification algorithm (see below for details) whose output provides detailed information about the location and extent of damage in patients with and without acquired language disorders. The lesion sites that are found to consistently result in very specific functional impairments (i.e. critical lesion sites) are then used to predict outcome in new patients and also offer unique insights into the functional anatomy of language. Those patients who perform unexpectedly well on a number of language tasks despite relatively focal damage to previously identified critical regions or relatively large left or right hemisphere strokes are invited to participate, upon obtaining informed consent, in a multifactorial fMRI paradigm designed to tease apart a whole range of language processing levels. The selection criteria for the fMRI component of the PLORAS project are both study-specific and constantly being updated as our knowledge about how very specific language impairments relate to lesion site progresses over time. For example, in the fMRI study by Seghier et al. (2014), the goal was to investigate how speech production can be supported after left putamen damage. Accordingly, patients with relatively focal left putamen damage were contacted to complete the PLORAS fMRI paradigm so as to identify brain areas with abnormally high activation during successful reading aloud and picture-naming responses. In short, the fMRI data allow us to explain inter-patient variability in lesion-deficit mappings by pinpointing which of the preserved brain regions are functioning normally or abnormally, and how they connect to one another. The fMRI

data can therefore be used to understand the neural pathways that support normal behaviour following brain damage. Neurologically-normal individuals are also invited to participate in the fMRI component of the project to define a range of normal responses and establish the degree to which each patient's fMRI pattern departs from that normal range. When possible, the behavioural and neuroimaging data are collected at multiple time points after stroke so as to sample the entire recovery process. By integrating cross-sectional and longitudinal behavioural, demographic and neuroimaging data, the main factors that determine long-term outcome and recovery after stroke start to emerge. This knowledge is then progressively translated into structure-function-recovery rules in the form of a clinical prognosis tool which is expected, in the future, to generate the most likely recovery trajectory for a new patient based on the site and extent of the stroke lesion, thereby informing clinical prognosis and therapy planning (see Figure 2.1).

2.1.1 My contribution to the PLORAS database

Over the course of the last 5 years, I have actively contributed (as part of a team effort) to the continuous expansion of the PLORAS database by participating in the recruitment, assessment and scanning of tens of stroke patients. I have also been involved in the collection and analysis of neuroimaging data from patients with brain tumours with the goal of better understanding how the gradual expansion of tumours versus the acute onset of strokes impacts upon the language system. Finally, I have set up a project in Chile (my mother country) which aims to enable, in the medium term, the realisation of studies investigating the effects of cultural (UK versus Chile) and linguistic (English versus Spanish) variables on language outcome and recovery after stroke. This has entailed translating the Comprehensive Aphasia Test (see below) into Spanish, sorting out the logistics of the project, securing ethical clearance, testing approximately 100 Chilean stroke survivors and obtaining their clinical scans (i.e. computerized tomography).

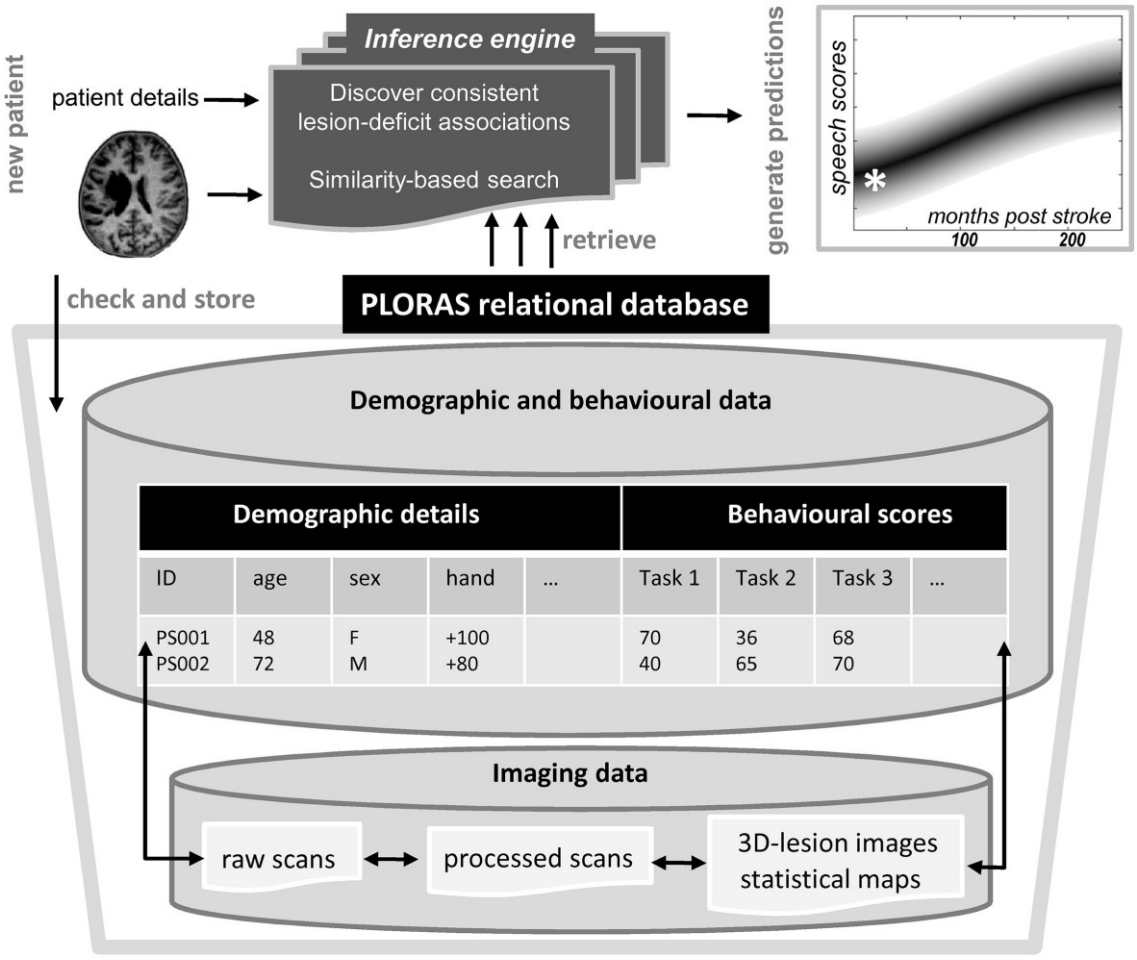

Figure 2.1: The PLORAS approach. The structural brain scan of a new patient is converted into a 3-dimensional image encoding the location and extent of damage incurred. The lesion image is then compared to those of all other patients in the database. The speech scores of the patients with similar lesions (both in terms of size and location) and demographic details are subsequently retrieved. By plotting the behavioural scores of the matching patients against time post-stroke, a prediction for the new patient is generated which indicates their likely time course of recovery with a confidence rating. Reproduced with permission (Elsevier) from Seghier et al. (2016).

2.2 The behavioural assessment tools

The speech and language abilities of all patients recruited to the PLORAS database are assessed with the Comprehensive Aphasia Test (CAT) (Swinburn et al., 2004). The CAT is a widely adopted and fully standardised test battery, which consists of a total of 27 different tasks. In comparison with other assessment tools, preference for the CAT is based on the following four main observations: (a) it has been shown to have robust psychometric properties as reflected by good validity and reliability measures (Swinburn et al., 2004); (b) the stimuli were designed to control for the most relevant

psycholinguistic variables such as word length, imageability, frequency as well as regularity (for further discussion, see Bruce and Edmundson, 2010; Howard et al., 2010a, b; Springer and Mantey, 2010); (c) it allows a comprehensive set of cognitive and language functions to be evaluated while being relatively quick to administer (a typical session lasts between 1 and 2 hours); and (d) it is currently being adapted into 12 different languages by the Collaboration of Aphasia Trialists to foster cross-linguistic comparative research on aphasia (Brady et al., 2014; Fyndanis et al., 2017).

As with any other language assessment tool, however, the CAT is not devoid of limitations (Bruce and Edmundson, 2010; Springer and Mantey, 2010). What these might be is contingent upon the purpose of assessment as well as the theoretical perspective of the tester. For example, whereas the main aim of a researcher may be to classify the patient's performance into one of the major aphasic syndromes, a clinician will likely be more interested in obtaining enough information about the affected and intact language abilities of the patient so as to set appropriate goals for therapy. Furthermore, the ecological validity of structured language assessment batteries is still a matter of debate. Another limitation could arise when subtests include too few items compromising the sensitivity of the language measures to capture milder deficits. In relation to the CAT, specifically, verb naming (i.e. task number 18 below) involves 5 items compared to the 24 used to evaluate object naming (i.e. task number 17 below), which poses constraints on the conclusions that can be drawn if one is particularly invested in studying verb processing deficits. Put differently, a subset of the tasks included in the CAT will provide an accurate description of moderate to severe behavioural impairments, but may need to be supplemented with other more taxing tests in some cases (or more sensitive measures such as reaction times). In particular, it is possible that patients who perform within the normal range on these tasks might nonetheless have milder impairments. With that being said, the CAT is a valid and reliable comprehensive battery that allows the patient's performance across a wide range of language tasks to be characterised in a relatively short time period. This summary profile of linguistic abilities can then be used

to pinpoint the overall severity and underlying nature of the observed language impairments.

With respect to the scoring system of the CAT, the authors encourage (for ease of comparison across tasks) the conversion (through a non-linear transformation) of raw scores into T-scores, which represent how well the patient performed relative to a reference population of 113 patients with aphasia, 56 of whom were tested more than once. For example, a T-score of 50 indicates the mean of the patient sample used to standardise the CAT, whereas a T-score of 60 represents one standard deviation above the mean. Most people without post-stroke aphasia would therefore be expected to score above the average of the patient standardisation sample on any given task from the CAT. The threshold for impairment is defined relative to a second reference population of 27 neurologically-normal controls. Specifically, it is the point below which the score would place the patient in the bottom 5% of the control population (Swinburn et al., 2004). Lower scores indicate poorer performance. Importantly, the two standardisation samples referred to before (i.e. 113 patients with aphasia and 27 neurologically-normal controls) are completely independent of the data being reported here (for more details on the standardisation samples, see Swinburn et al., 2004).

As stated in the CAT manual (p. 71), the main advantages of converting raw scores into T-scores is that this allows: (i) scores from different tasks to be compared because they have been put on a common scale; and (ii) the use of parametric statistics given that T-scores are normally distributed scores with a mean of 50 and a standard deviation of 10.

The task stimuli, administration and scoring system were taken from the original English version of the CAT (Swinburn et al., 2004) and occurs in the following order:

(1) **Line bisection:** this task visually presents three horizontal lines on a sheet of paper with instructions to draw a vertical line through the midpoint of each horizontal line. There is a practice trial at the beginning where the participant can be given as many demonstrations as required to ensure comprehension of the

task. A template is provided to estimate the degree of deviance from the line centre with three marks along the horizontal axis indicating normal, mild and severe visuospatial defects. Vertical lines placed at the centre score 0; between the centre (0) and the first (0.5) mark score 0.5; between the first (0.5) and second (1.0) mark score 1.0; and between the second (1.0) and third (2.0) mark score 2.0. A positive or negative sign is attached to the score depending on whether the deviation from the centre is towards the right (+) or towards the left (-). The scores are then added up to yield a total score ranging from 0 to ±6. A total T-score equal to or below 39 signals impaired performance.

(2) **Semantic associations:** this task visually presents five pictures of objects simultaneously. The instructions are to match the picture at the centre (e.g., mitten) with one of four possible alternatives according to the strongest semantic association (e.g., hand, sock, jersey, and lighthouse). The inclusion of a semantically related distractor (e.g., sock) encourages deeper levels of semantic processing/control. There are a total of 10 test trials plus a practice one at the beginning. Correct responses are given a score of 1; incorrect responses are given a score of 0. A total T-score equal to or below 47 signals impaired performance.

(3) **Word fluency:** this task prompts the participant to say out loud as many words as possible within a given semantic category (semantic fluency) and starting with a given letter (letter fluency) in a period of 1 minute each. There is a practice trial for the semantic and letter fluency components of the task involving the semantic category “clothes” and the letter “b” followed by the test trials “animals” and “s”, respectively. Each correct word attracts a score of 1. A total T-score equal to or below 57 signals impaired performance.

(4) **Recognition memory:** this task visually presents each of the ten central items from the CAT semantic associations task (one at a time) along with three unrelated distractors in a 2×2 array. The instructions are to indicate which one of the four pictures on display had been seen before. There are a total of ten test trials plus a practice one at the beginning. Correct responses are given a score of 1; incorrect

responses are given a score of 0. A total T-score equal to or below 43 signals impaired performance.

(5) **Gesture object use:** this task visually presents six pictures of common objects (e.g., scissors) with instructions to demonstrate how each object would be used. There is a practice trial at the beginning to ensure comprehension of the task. Correct responses are given a score of 2; incorrect responses are given a score of 0. If the action or the orientation of the gesture is incorrect or if a body part is used as the object itself, a score of 1 is granted. A total T-score equal to or below 51 signals impaired performance.

(6) **Arithmetic:** this task requires the participant to solve six simple arithmetic problems including subtraction, addition and multiplication. The instructions are to point to the correct answer from a choice of five. Correct responses are given a score of 1; incorrect responses are given a score of 0. A total T-score equal to or below 40 signals impaired performance.

(7) **Auditory word-to-picture matching:** this task involves hearing a word and selecting the picture, among four possible alternatives (2×2 array), that best matches the meaning of the heard word. Apart from the target, each trial includes one semantic, one phonological and one unrelated distractor. There are a total of fifteen test trials plus a practice one at the beginning. Immediate correct responses are given a score of 2; incorrect responses are given a score of 0; correct responses after repetition of the target upon request, self-correction or delay (> 5 seconds) are given a score of 1. The task is discontinued if the participant makes 4 consecutive incorrect responses. A total T-score equal to or below 51 signals impaired performance.

(8) **Visual word-to-picture matching**: this task involves a written word at the centre of the page surrounded by four possible pictures (2×2 array). The participant has to select the picture that best matches the meaning of the written word. As in the auditory word-to-picture matching, semantic, phonological and unrelated distractors are part of each trial. There are a total of fifteen test trials plus a practice

one at the beginning. Immediate correct responses are given a score of 2; incorrect responses are given a score of 0; correct responses after self-correction or delay (> 5 seconds) are given a score of 1. The task is discontinued if the participant makes 4 consecutive incorrect responses. A total T-score equal to or below 53 signals impaired performance.

(9) **Auditory sentence-to-picture matching:** this task involves hearing a sentence and selecting the picture, among four possible alternatives (2×2 array), that best matched the meaning of the sentence. There are a total of sixteen test trials plus a practice one at the beginning. The task spans a wide range of syntactic structures such as reversible, active, passive and embedded sentences. The scoring system for this task is identical to that for the auditory word-to-picture matching task. The task is discontinued if the participant makes 4 consecutive incorrect responses. A total T-score equal to or below 60 signals impaired performance.

(10) **Visual sentence-to-picture matching:** this task involves a written sentence at the centre of the page surrounded by four possible pictures. The participant has to select the picture that best matches the meaning of the written sentence. As in the auditory sentence-to-picture matching task, a range of syntactic structures such as reversible, active, passive and embedded sentences are assessed. There are a total of sixteen test trials plus a practice one at the beginning. The scoring system for this task is identical to that for the visual word-to-picture matching task. The task is discontinued if the participant makes 4 consecutive incorrect responses. A total T-score equal to or below 57 signals impaired performance.

(11) **Auditory comprehension of paragraphs:** this task involves listening to two short stories that are followed by four “Yes/No” questions each. There are two different wordings of the same question, with a score of 1 being given only if the two versions are associated with correct responses (i.e. the maximum score for this task is 4). If only one version is correctly answered, a score of 0 is allocated. A total T-score equal to or below 43 signals impaired performance.

(12) **Word repetition:** this task aurally presents words, one at a time, with instructions to repeat them out loud. Psycholinguistic properties of the stimuli such as imageability, frequency and length are controlled throughout. There are a total of sixteen test trials plus a practice one at the beginning. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target word are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the auditory word-to-picture matching task. A total T-score equal to or below 56 signals impaired performance.

(13) **Complex word repetition:** this task aurally presents three morphologically complex words (e.g., unthinkable), one at a time, with instructions to repeat them aloud. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target word are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the auditory word-to-picture matching task. A total T-score equal to or below 55 signals impaired performance.

(14) **Non-word repetition:** this task aurally presents five nonsense words (e.g., gart) of increasing length and phonological complexity, one at a time, with instructions to repeat them out loud. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the auditory word-to-picture matching task. A total T-score equal to or below 51 signals impaired performance.

(15) **Digit span:** this task involves hearing and repeating digit strings grouped in six progressive levels of length. At each level, the first of a pair of digit strings is read out. If the response is correct, the next level is reached; whereas in the opposite scenario a second opportunity is granted by presenting the second digit string. The task starts with two digits and builds up to seven digits. Phonemic and apraxic errors are not penalised. Articulatory errors (e.g., dysarthric distortions) not

affecting the perceptual identity of the target word are also accepted. The total score is obtained by multiplying the number of digits in the digit string of maximum length successfully repeated by two. A total T-score equal to or below 50 signals impaired performance.

(16) **Sentence repetition:** this task involves hearing and repeating sentences grouped in four progressive levels of length. At each level, the first of a pair of sentences is read out. If the response is correct, the next level is reached; whereas in the opposite scenario a second opportunity is granted by presenting the second sentence. The task varies from sentences with three content words at the lower level (e.g., The cat chased the bird) up to six content words at the upper level (e.g., The boy and girl climbed the hill and admired the view). Phonemic and apraxic errors are not penalised. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target word are also accepted. The total score is obtained by multiplying the number of content words from the longest sentence successfully repeated by two. A total T-score equal to or below 56 signals impaired performance.

(17) **Spoken picture naming:** this task visually presents line drawing pictures of objects (e.g., knife), one at a time, with instructions to name them aloud. Psycholinguistic properties of the stimuli such as imageability, frequency, animacy and length are controlled throughout. There are a total of twenty-four test trials plus a practice one at the beginning. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the visual word-to-picture matching task. The task is discontinued if the participant makes 8 consecutive incorrect responses. A total T-score equal to or below 61 signals impaired performance.

(18) **Action naming:** this task visually presents line drawing pictures, one at a time, depicting high-frequency actions (e.g., typing) with instructions to say what the

person in each picture is doing. There are a total of five test trials plus a practice one at the beginning. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the visual word-to-picture matching task. A total T-score equal to or below 59 signals impaired performance.

(19) **Spoken picture description:** this task is designed to provide the means of obtaining a sample of connected speech with a reliable scoring system. The participant is shown a picture depicting a complex scene and asked to verbally describe what is happening for 1 minute. The speech sample is then rated based on the total number of appropriate information carrying words (i.e. words that convey any information as opposed to content words) minus the total number of inappropriate information carrying words (i.e. information carrying words that are mistakenly selected from the mental lexicon), plus syntactic variety (on a 0-6 scale), grammatical well-formedness (on a 0-6 scale) and speed of speech production (on a 0-3 scale). A total T-score equal to or below 60 signals impaired performance.

(20) **Word reading:** this task visually presents words, one at a time, with instructions to read them aloud. Psycholinguistic properties of the stimuli such as frequency, length and spelling-to-sound correspondence are controlled throughout. There are a total of twenty-four test trials plus a practice one at the beginning. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target word are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the visual word-to-picture matching task. A total T-score equal to or below 61 signals impaired performance.

(21) **Complex word reading:** this task visually presents three morphologically complex words (e.g., informative), one at a time, with instructions to read them aloud. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity

of the target word are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the visual word-to-picture matching task. A total T-score equal to or below 57 signals impaired performance.

(22) **Function word reading:** this task visually presents three function words (e.g., but), one at a time, with instructions to read them aloud. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target word are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the visual word-to-picture matching task. A total T-score equal to or below 48 signals impaired performance.

(23) **Non-word reading:** this task visually presents five nonsense words (e.g., fask), one at a time, with instructions to repeat them out loud. Articulatory errors (e.g., dysarthric distortions) not affecting the perceptual identity of the target are scored as correct responses. Verbal, phonemic, neologistic and apraxic errors are scored as incorrect responses. The scoring system for this task is identical to that for the visual word-to-picture matching task. A total T-score equal to or below 56 signals impaired performance.

(24) **Copying text:** this task visually presents letters and words that the participant is prompted to copy: five letters from upper to upper case and five letters from lower to upper case. Additionally, the patient is asked to copy three words (in lower case) using only capital letters. Each correct letter is given a score of 1; incorrect letters are given a score of 0. A total T-score equal to or below 50 signals impaired performance.

(25) **Written picture naming:** this task visually presents pictures of objects, one at a time, with instructions to write their names down. Psycholinguistic properties of the stimuli such as frequency, regularity and length are controlled throughout. There are a total of five test trials plus a practice one at the beginning. Letters in the correct position are given a score of 1 each. Substitutions, omissions and

transpositions are given a score of 0. One point is deducted from the total score if one or more letters are added to the target word. Semantically related verbal paraphasias are scored as incorrect responses. A total T-score equal to or below 54 signals impaired performance.

(26) **Writing-to-dictation:** this task involves hearing words (one at a time) and writing them down as accurately as possible. There are a total of five test trials plus a practice one at the beginning. The test items comprise a concrete word “man”, an irregular concrete word “yacht”, an abstract word “idea”, a morphologically complex word “undrinkable”, and a non-word “blosh”. The scoring system for this task is identical to that for the written picture naming task. A total T-score equal to or below 57 signals impaired performance.

(27) **Written picture description:** this task presents the participant with the same picture used in the spoken picture description task. The instructions are to write down what is happening in the picture for 3 minutes. The sample of connected writing is then rated based on the total number of appropriate information carrying words (i.e. words that convey any information as opposed to content words) minus the total number of inappropriate information carrying words (i.e. information carrying words that are mistakenly selected from the mental lexicon), plus grammatical well-formedness (on a 0-6 scale). A total T-score equal to or below 65 signals impaired performance.

2.3 Phonological and semantic decisions

A subset of the patients from the PLORAS database completed, in addition to the standard CAT assessment, the phonological and semantic decision tasks that previous studies used to dissociate phonological and semantic processes in the left inferior frontal and inferior parietal cortices with functional magnetic resonance imaging (e.g., Devlin et al., 2003; Gitelman et al., 2005) and transcranial magnetic stimulation (e.g., Gough et al., 2005; Sliwinska et al., 2015). These data were acquired in the context of three previous research projects conducted by three former group members (Varun Sethi,

Abigail Boulton and Eleanor Paine) who investigated the effect of right- and left-hemisphere stroke lesions on phonological and semantic decisions. Patients from the PLORAS database who met the following selection criteria were invited to take part in these studies: (a) aged over 18; (b) no other neurological or major psychiatric disorder; (c) right-handed (pre-morbidly); (d) English as first language (either monolingual or bilingual); (e) normal or corrected-to-normal vision and hearing (as per self-reports); (f) left-hemisphere or right-hemisphere lesion (as attested by a clinical neurologist); (g) no or minimal damage in the contralateral hemisphere according to a clinical neurologist. For patient details, see Chapter 3.

Specifically, the speed and accuracy of the patients' responses to two tasks that required either making judgements about the sound structure (e.g., do two words sound the same?) or meaning (e.g., are two words related in meaning?) of visually presented pairs of words were obtained. The word stimuli were provided by Professor Joseph T. Devlin, who is the senior author on the TMS studies by Gough et al. (2005) and Sliwinska et al. (2015). The complete list of stimuli is presented in Appendix 1. The experiment was carried out in a well-lit and quiet room to minimise any external distractors; and all patients were exposed to the same set of stimuli and in the same order. Given the possibility that switching costs may be greater for patients than neurologically-normal controls, the order of the tasks was also held constant across participants. The motivation for this decision was to allow behavioural performance to be compared between patients and neurologically-normal controls, while minimising the contribution of confounding sources of variance. Stimulus presentation was achieved via a patient-dedicated laptop computer operating under Microsoft Windows XP Professional. Cogent 2000 (http://www.vislab.ucl.ac.uk/cogent.php), running under MATLAB 7, was utilised to design and display the experimental stimuli. Responses were collected via a two-button response pad (Current Designs, Philadelphia, USA) plugged into the testing laptop. Reaction times were measured from stimulus onset to the point where the participant's response begun.

Prior to data acquisition, a practice run involving 12 different stimulus pairs was conducted to ensure that (i) all equipment were functioning normally, (ii) the patient's responses were being recorded, (iii) the task had been correctly understood and (iv) the patient could see the stimuli well. The items that were displayed in the practise run were discarded and not reused in the main experiment (for more details, see Appendix 1). Across the phonological and semantic decision tasks, word stimuli were matched for concreteness, familiarity, written word frequency, number of letters and number of syllables. Concreteness and familiarity ratings were taken from the MRC psycholinguistic database (Coltheart, 1981), while British English word frequencies were obtained from the CELEX lexical database (Baayen et al., 1995). Each task consisted of 2 blocks and each block comprised 32 unique word pairs counter-balanced for Yes-No responses. Within each block, half the word pairs were associated with "Yes" (same/related) responses and half with "No" (different/unrelated) responses, with the order of "Yes/No" responses randomised. All visual stimuli were presented in black Arial font (size 32) on a white background for a maximum duration of 3000 ms including a 200 ms fixation cross and a 400 ms inter-trial interval. More than one block was administered in case the patient's errors could be attributed to difficulties understanding the task or other external factors. At the beginning of each task, the patients were clearly instructed to make a decision as quickly and accurately as possible, by pressing the "Yes" or "No" buttons on the response pad. For scoring purposes, the first block was always preferred unless the total number of correct responses achieved on the second block exceeded that on the first block by at least two trials. Performance on the phonological and semantic decision tasks was expressed in terms of a composite measure (i.e. "efficiency") that combined the speed and accuracy of the patients' responses into a single summary score as per the following calculation: (percent accuracy ÷ median correct reaction time) × 1000.

2.4 Basic principles of MRI physics

The material presented in this section has been taken from the excellent introductory articles by Pooley (2005), Bitar et al. (2006), Plewes et al. (2012) and Currie et al. (2013).

Magnetic resonance imaging (MRI) is a non-invasive technique which uses a strong magnetic field to generate images of the brain. The key components of the MRI scanner are: (i) the main electromagnet; (ii) the radiofrequency coils; (iii) the gradient coils; and (iv) the shimming coils (see Figure 2.2). The main electromagnet is made of a superconducting metal-alloy immersed in cryogenic liquid helium to produce a high-strength magnetic field with little heat disposition. The main magnetic field (B_0) of an MRI system originates from a large electric current flowing through wires that are formed into a loop in the magnet. Most research MRI systems operate at 1.5 Tesla or 3 Tesla, which indicates the strength of the generated static magnetic field. For example, 1.5 Tesla (T) is equivalent to approximately 30,000 times the earth's magnetic field. The radiofrequency (RF) coils generate electromagnetic energy in the megahertz range (B_1) and can be thought of as a very basic transmitting and receiving antenna. For neuroimaging, a separate RF receiver coil is placed around the participant's head to enhance the detection of the emitted MRI signals back from the brain. The gradient coils are inserted in the MRI system and are used to slightly distort the main magnetic field in a very precise manner along the x (right-left), y (anterior-posterior) or z (superior-inferior) directions. The shimming coils are used to correct local susceptibility artefacts and inhomogeneities in the main B_0 field.

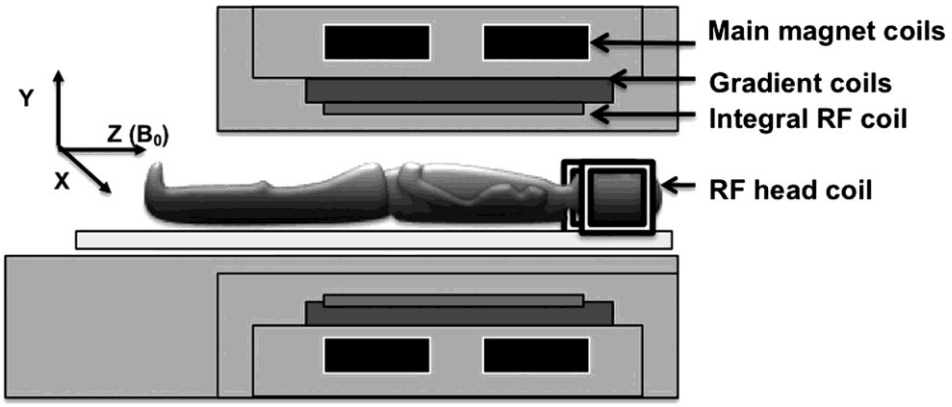

Figure 2.2: MRI scanner. The figure serves as a graphical illustration of the main components of a standard MRI system. Reproduced with permission (BMJ Publishing Group Ltd.) from Currie et al. (2013).

Hydrogen nuclei (associated with fat and water molecules) are abundant in the human body including the brain and comprise a single proton. Hydrogen protons can be portrayed as positively charged spheres that are constantly spinning around their axes, which gives rise to a magnetic field (i.e. they behave similar to small magnets with charge and mass). Under normal conditions, magnetic moments cancel each other out because they are typically randomly oriented. However, when placed in a strong magnetic field (B_0), protons tend to align with the external field because it is the state that requires the least energy. This results in a sum magnetisation (M_0) parallel to the applied magnetic field (see Figure 2.3A). Given that all protons are exposed to the same magnetic field, they move in a particular way referred to as precession (i.e. like a spinning top). The net magnetisation (M_0) cannot be measure because it is aligned with the external magnetic field (B_0) in the z (or foot-to-head) direction. Therefore, RF pulses that match the precession frequency of the population of spins (aka Larmor frequency) are switched on and off to flip over the nuclei so that they fall out of alignment with B_0. In other words, when exposed to the B_1 field (i.e. the RF pulses), protons are forced to move in the same direction and at the same time (i.e. in phase) resulting in a new "transverse" component of the magnetisation. In contrast to the ''longitudinal'' component which is aligned with (i.e. parallel to) the B_0 field, the transverse magnetisation lies in the x-y plane (perpendicular to the B_0 field) where it is free to precess. Importantly, the new rotating magnetisation vector is able to induce a signal in the receiver coil because as soon as

the RF pulse is switched off the protons start to return to a lower energy state (i.e. towards the longitudinal magnetisation vector) falling out of alignment with each other (i.e. dephasing), which is known as relaxation. Relaxation takes the following two forms: the magnetisation starts to grow back in the longitudinal direction (aka longitudinal or T1 relaxation/recovery) and away from the transverse plane (aka transverse or T2 relaxation/decay). Critically, the rate at which these processes occur is tissue-specific (see Figure 2.3B).

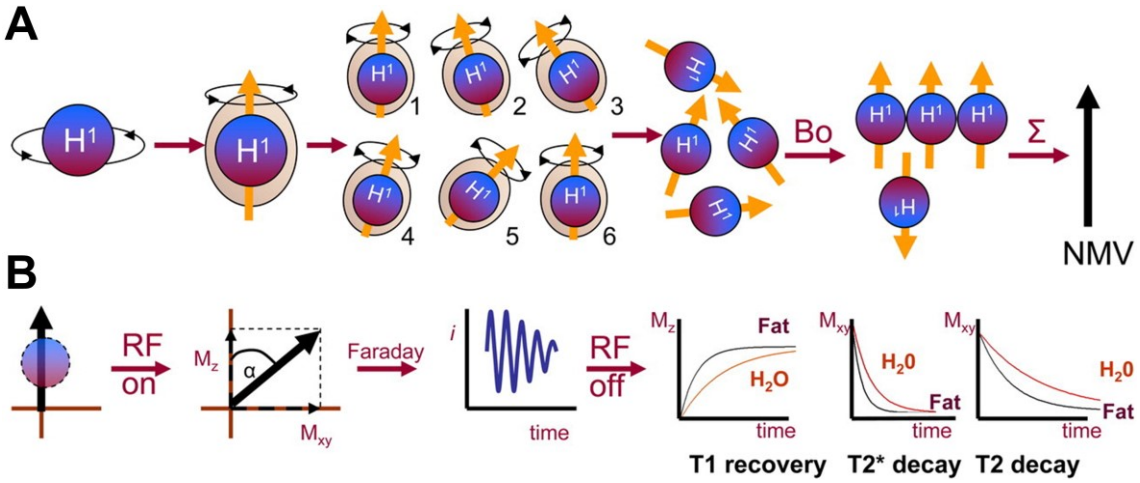

Figure 2.3: MRI signal generation process. (A) Spinning hydrogen (H^1) protons generate magnetic fields (tan) whose direction can be depicted as vectors (yellow). In the absence of an external magnetic field, the protons are randomly oriented with respect to each other cancelling their collective magnetic effect. However, when a strong external magnetic field (B_0) is applied, more protons align with B_0 than against it resulting in a net magnetisation vector (NMV). (B) The bulk magnetic effect cannot be measured by the MRI scanner because the receiver coil is only sufficiently sensitive to time-varying magnetic fields. In order to induce a signal, an RF pulse is delivered to tip the spinning protons from their alignment with B_0 so that they begin precessing about the external magnetic field. As a consequence, a new transverse magnetisation component (M_{xy}), as opposed to the longitudinal magnetisation component (M_z), is generated which by Faraday's law induces a current in the receiver coil. When the RF pulse is turned off, the protons start dephasing and T1 recovery and T2 and T2* decay occur. Reproduced with permission (Radiological Society of North America) from Bitar et al. (2006).

The rate of regrowth of longitudinal magnetisation for each tissue type (i.e. white matter, grey matter and cerebrospinal fluid) is a fundamental source of contrast in T1-weighted (structural) brain images, with T1 being a time constant that indicates how long it takes for longitudinal magnetisation to reach approximately 63% of its final value. Put differently, given that T1 and T2 relaxation occur simultaneously but independently (i.e.

their time constants are not the same) and that the recovery of longitudinal magnetisation differs across tissue types, a brain image can be acquired at a time when these T1 curves are widely separated so as to differentiate one tissue from another (i.e. T1-weighted image). Loss of phase coherence across a population of protons (after the emission of the RF pulse) has, on the other hand, two main sources: (i) spin-spin interactions within the tissues (i.e. T2 relaxation) and (ii) inhomogeneity within B_0 due to changes in tissue magnetic ‘‘susceptibility’’ (e.g., paramagnetic tissues enhance the magnetic field and diamagnetic tissues weaken the magnetic field). Crucially, T2* relaxation is a time constant that describes the actual rate of decay of the signal (i.e. transverse magnetisation) resulting from the combined effect of spin-spin interactions and inhomogeneities within the local static magnetic field, which basically means that T2 relaxation and T2* relaxation can be modelled separately (see Figure 2.3B). By choosing appropriate imaging parameters, differences in signal intensity between tissue types will be weighted according to their T1, T2 or T2* relaxation times. For instance, the difference between a T1-weighted and T2*-weighted scan resides primarily in the choice of repetition (TR) and echo times (TE), which govern the time interval between successive RF excitation pulses and between excitation and data acquisition, respectively. Of note, T2*-weighted images are sensitive to the amount of deoxygenated haemoglobin present and form the contrast basis for functional MRI, whereas T1-weighted images form the contrast basis for structural MRI.

Through the use of spatially varying magnetic field gradients (i.e. gradient pulses), it is possible to determine the location of the signal in order to create MRI images. In more detail, an axial slice of the brain (of a certain thickness) is selectively targeted by applying a magnetic field gradient at a specific z location (i.e. slice-selection gradient) that modulates the precessional frequency of protons; immediately followed by a band-limited RF excitation pulse matching the Larmor frequency range of the spins in that slice (a phenomenon called resonance), which flips the net magnetisation from the longitudinal axis over into the transverse plane. Subsequently, a phase-encoding

gradient and a frequency-encoding gradient (the latter perpendicular to the former) are used to acquire information about the distribution of the protons within the slice itself: they selectively modify the angle and frequency of precession at specific locations along the y and x directions. Together, magnetic field gradients in MRI act to change the actual strength of the main static magnetic field over space allowing for the sampling and localisation of the "signals" from protons at the level of very small 3-dimensional volumetric units (i.e. voxels), via the formation of image sections and pixels in those sections (see Figure 2.4).

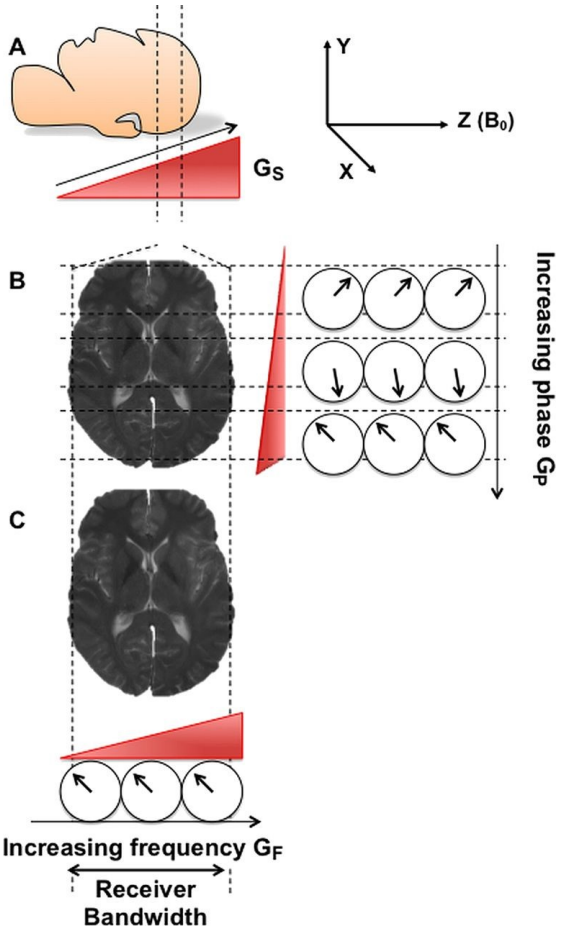

Figure 2.4: MRI signal location and encoding. Through the application of gradients, linear variations of the magnetic field strength are introduced. (A) The slice-selection gradient (G_S) changes the precessional frequency of protons at specific locations along the z axis (parallel to the main magnetic field), which (in conjunction with a tailored RF pulse) determines the section of the brain to be imaged. (B) The phase-encoding gradient (G_P) induces a phase shift in the spinning protons so that the phase of the spin can be detected and encoded. (C) The frequency-encoding gradient (G_F) induces a shift in the precessional frequency so that the location of the spinning nuclei can be detected. Together, the information gathered after applying the magnetic field gradients allows the MRI system processor to compute the exact location and amplitude of the signal. Reproduced with permission (BMJ Publishing Group Ltd.) from Currie et al. (2013).

Brain images that are sensitive to specific properties of the underlying tissue can be collected by manipulating the angle of the radiofrequency excitation pulses (or flip angle) and the timing of the gradient pulses along with the imaging parameters to assemble scanning sequences that target one of the previously described tissue weightings. Reconstruction of the recorded MRI signal into an image is achieved via its digitalisation and storage in matrix format (called “k-space”) in the central computer. Data points in k-space, which encode the spatial frequency information of the MRI signal, are converted into an image using a computationally efficient mathematical process known as a Fourier transform.

2.5 Structural MRI data acquisition, pre-processing and lesion identification

Out of a total of 288 patients reported in the current thesis, 270 underwent structural MRI at the UCL Wellcome Centre for Human Neuroimaging. The remaining 18 patients were scanned at the Birkbeck-UCL Centre for Neuroimaging (on a 1.5 T Avanto scanner); for patient details, see experimental chapters. Four different MRI scanners (Siemens Healthcare, Erlangen, Germany) were used to acquire the T1-weighted structural images: 150 patients were imaged on a 1.5 T Sonata scanner, 115 on a 3 T Trio scanner, 18 on a 1.5 T Avanto scanner, and 5 on a 3 T Allegra scanner. For anatomical images acquired on the 1.5 T Avanto scanner, a 3D magnetisation-prepared rapid acquisition gradient-echo (Mugler and Brookeman, 1990) sequence was used to acquire 176 sagittal slices with a matrix size of 256 × 224, yielding a final spatial resolution of 1 mm isotropic voxels (repetition time/echo time/inversion time = 2730/3.57/1000 ms). For anatomical images acquired on the other three scanners, an optimised 3D modified driven equilibrium Fourier transform (Deichmann et al., 2004) sequence was used to acquire 176 sagittal slices with a matrix size of 256 × 224, yielding a final spatial resolution of 1 mm isotropic voxels: repetition time/echo time/inversion time = 12.24/3.56/530 ms and 7.92/2.48/910 ms at 1.5 T and 3 T, respectively. The T1-weighted image for each patient was then submitted to a fully automated lesion identification procedure for lesion detection and delineation (see below for details).

Critically, this converts a scanner-sensitive raw image into a quantitative assessment of structural abnormality that is independent of the scanner used (because voxel intensities are normalised with respect to those observed in neurologically-normal controls imaged on the same scanners). Moreover, the quality of the generated lesion images is evaluated by visually inspecting the results. Three types of lesion identification errors, which might differ from manually drawn lesions, have been detected. First, the lesion extent includes cerebrospinal fluid in enlarged ventricles. Second, cortical atrophy (e.g., around the dorsal parietal lobes) can sometimes be included in the lesion image. Third, the automated approach can miss small cortical lesions where there is normal inter-subject variability in sulci. In addition, there are potential errors that arise in both automated and manually defined lesions, particularly in the specification of the border of the lesion which is typically gradual rather than categorical. I did not attempt to correct any of these errors (none of the lesions were manually drawn) and they therefore increased “noise” in the analyses, which may result in false negatives but not false positives.

Normally, the manual segmentation of abnormal brain tissue by an expert neurologist or neuro-radiologist is assumed to be the gold-standard method for lesion identification (Wilke et al., 2011). However, the manual tracing of lesions is time-consuming, laborious and operator-dependent (Fiez et al., 2000). To overcome some of these limitations and given the scale of the PLORAS project, the T1-weighted anatomical whole-brain volume of each patient was pre-processed in SPM8 (Wellcome Centre for Human Neuroimaging, University College London, London, UK) running under MATLAB (MathWorks, Natick, Massachusetts, USA) with an automated lesion identification toolbox using default parameters (Seghier et al., 2008a). The procedure combines an optimized segmentation-normalisation routine (Ashburner and Friston, 2005; Crinion et al., 2007) with an outlier detection algorithm according to the fuzzy logic clustering principle (Seghier et al., 2007). The outlier detection algorithm assumes that a lesioned brain is an outlier in relation to normal (control) brains. The output includes two 3D lesion

images in standard Montreal Neurological Institute (MNI) space, generated at a spatial resolution of 2 × 2 × 2 mm^3. The first is a fuzzy lesion image that encodes the degree of structural abnormality on a continuous scale from 0 (completely normal) to 1 (completely abnormal) at each given voxel relative to normative data drawn from a sample of 64 neurologically-normal controls. A voxel with a high degree of abnormality (i.e. a value near to 1 in the fuzzy lesion image) therefore means that its intensity in the segmented grey and white matter deviated markedly from the normal range. The second is a binary lesion image, which is simply a thresholded (i.e. lesion/no lesion) version of the fuzzy lesion image.

In more detail, the lesion identification procedure outlined above incorporates an "extra" tissue class (in addition to the normal grey matter, white matter and cerebrospinal fluid tissue probability maps) and fuzzy clustering with fixed-prototypes to detect outlier lesioned voxels across the whole brain. The following 4 steps are involved (see Figure 2.5):

I. **Segmentation and normalization:** the segmentation and spatial normalisation of T1-weighted structural brain scans into a common reference space is achieved via a unified segmentation-normalisation scheme which combines image registration, tissue classification and bias correction within a single iterative probabilistic model. The spatial normalisation component entails establishing a one-to-one mapping between the brains of individual patients by registering each brain to an average brain reference template (typically in MNI space; Mazziotta et al., 2001b) through the application of a set of linear and non-linear spatial transformations to the raw T1-weighted image. This is particularly important in the context of group-level analyses because of the known inter-subject variability in brain size and shape which would render basic inferential assumptions such as voxel X in subject 1 is roughly the same as voxel X in subject 2 (and so on and so forth) invalid. The segmentation component, on the other hand, allows the compartmentalisation of each individual brain into grey matter, white matter and cerebrospinal fluid tissue classes. A bias field correction

component is used to account for magnetic field inhomogeneities during image acquisition. Critically, it has been shown that the combination of the spatial normalisation, segmentation and bias correction components above in the inversion of a single unified model yields better results than serial applications of each of these modules for both undamaged (Ashburner and Friston, 2005) and damaged (Crinion et al., 2007) brains. The current implementation slightly modifies the tissue segmentation component by adding an "extra" tissue class to account for the presence of "atypical" voxels within the lesion that do not match the expected tissue types; namely, grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF). In other words, the misclassification of damaged tissue as intact GM or WM is prevented by allowing voxels with abnormal signal intensities (i.e. the lesion) relative to the range of voxel values in the average GM and WM probability maps derived from the International Consortium for Brain Mapping (Mazziotta et al., 2001a) to be included in the "extra" tissue class. Importantly, these tissue probability maps encode how likely it is that any one voxel in the brain belongs to the GM, WM or CSF class.

To explicitly model unexpected/abnormal voxels (i.e. those that show a mismatch between the spatial priors and the expected T1 signal), an "extra" tissue class is initially created by averaging the WM and CSF prior probabilities and further refined as part of an iterative process in which the estimated extra class is used as a prior for the next segmentation run. The addition of an "extra" tissue class to the segmentation component basically ensures that abnormal voxels are given a low probability of being intact GM or WM. The output of the modified segmentation-normalisation step, therefore, comprises four normalised and segmented tissue classes per subject (i.e. GM, WM, CSF and the extra class).

II. **Spatial smoothing:** the smoothing of brain images is a critical step in that it ensures that any residual differences in brain anatomy across subjects are supressed. The normalised GM and WM segments are, therefore, spatially smoothed with an 8 mm full width at half maximum Gaussian kernel to minimise fine-scale inter-subject anatomical variability.

III. **Outlier detection:** fuzzy clustering with fixed prototypes (for more details, see Seghier et al., 2007) is used to identify outlier voxels across the GM and WM segments as compared with normative data drawn from a sample of 64 neurologically-normal individuals. As a result, two fuzzy sets are generated: one for GM and another for WM, that encode (on a scale from 0 to 1) the degree to which each voxel probability of being either GM or WM departs from typical GM or WM voxel values. In other words, it detects outlier voxels in the normalised and spatially smoothed GM and WM segments whose probability of being either GM or WM is very low relative to control GM and WM segments that have been pre-processed in the exact same way.

IV. **Lesion definition (grouping):** in contrast to the subjective assessment of lesioned tissue by trained specialists, the WM and GM fuzzy sets are combined to form the fuzzy set union (or fuzzy lesion image) which provides an objective quantification of the degree of structural abnormality at each and every voxel of the brain. This fuzzy lesion image can subsequently be thresholded at a certain value (typically 0.3) to generate a binary lesion image that indexes the presence (i.e. value of 1) or absence (i.e. value of 0) of a lesion on a voxel-by-voxel basis at a spatial resolution of 2 × 2 × 2 mm^3. The binary lesion images are used to delineate the lesions, to estimate lesion size, to measure the degree of damage to specific regions of interest, and to create lesion overlap maps (e.g., Frank et al., 1997).

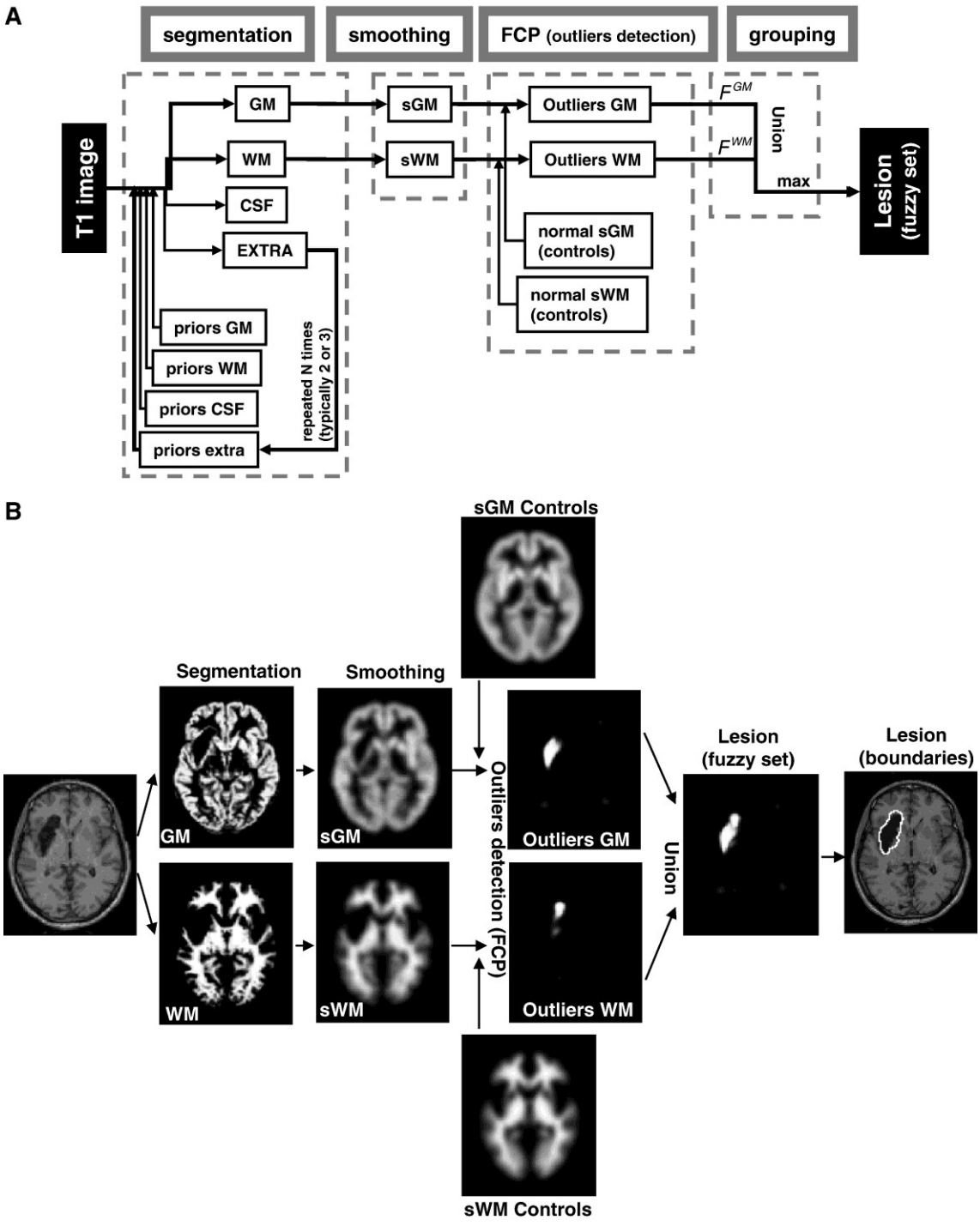

Figure 2.5: Automated lesion identification. (A) All the steps that form part of the automated lesion identification procedure used to define and delineate the stroke patients' lesions from T1-weighted brain images are depicted (see main text for details). (B) The output of each step is illustrated with a real example. Reproduced with permission (Elsevier) from Seghier et al. (2008a).

2.6 Neurophysiological basis of the fMRI BOLD signal

Functional magnetic resonance imaging (fMRI) has rapidly become the most popular tool to probe whole-brain function non-invasively since its inception. The fMRI signal, originally described by Ogawa et al. (1990a, b), is an indirect measure of neural activity that capitalizes on hemodynamic changes resulting from local increases in oxygenated blood flow to regions of the brain where metabolic demands are high (relative to baseline levels) due to the presence of active neurons (Arthurs and Boniface, 2002; Nair, 2005; Logothetis, 2008). The correct interpretation of the blood-oxygenation-level-dependent (BOLD) response therefore heavily rests on a thorough understanding of the mechanisms by which the underlying neural activity give rise to its hemodynamic correlate (i.e. the so called neurovascular coupling); see Figure 2.6. In two landmark studies, for instance, Logothethis and colleagues acquired simultaneous recordings of local field potentials (LFPs) and hemodynamic responses from anaesthetized (Logothetis et al., 2001) and awake (Goense and Logothetis, 2008) monkeys to show that the BOLD response is a surrogate measure of input and intra-cortical processing rather than pyramidal cell output activity. Critically, the authors also demonstrated that LFPs are better predictors of the BOLD response than multiple-unit or single-unit spiking activity (but see Heeger and Ress, 2002 and Mukamel et al., 2005). Put simply, the BOLD response reflects neuronal mass activity over millions of neurons at each voxel.

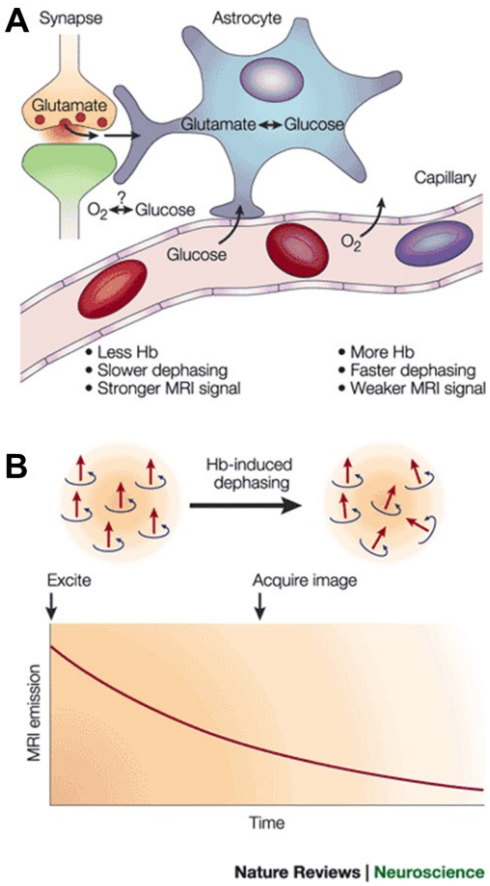

Figure 2.6: The fMRI BOLD signal. (A) The proposed relationship between neural activity, metabolic demands and hemodynamic changes (i.e. neurovascular coupling). (B) Paramagnetic substances such as deoxy-haemoglobin act as endogenous contrast agents by inducing inhomogeneities within the local magnetic field resulting in a faster dephasing of spinning protons and thus more rapid decay of transverse magnetisation (i.e. a shorter T2* relaxation). In other words, changes in the ratio of oxygenated to deoxygenated blood lead to differences in T2* relaxation times, which can be used to dissociate active brain regions from non-active ones. Reproduced with permission (Springer Nature) from Heeger & Ress (2002).

The appeal of the BOLD signal stems from the fact that changes in deoxy-haemoglobin concentration act as an endogenous paramagnetic contrast agent that can be detected by the MRI scanner (Kim and Ogawa, 2012). Neurobiologically, it is thought that neural activity increases in response to task demands. This causes astrocytes and neuronal cells to send vasodilatory signals into nearby arterioles and capillaries, which in turn lead to focal increases in cerebral blood flow. The energy demands (glucose and oxygen) of the active neurons can therefore be met. Critically, however, the increase in cerebral blood flow surpasses the cerebral metabolic rate of oxygen consumption resulting in an imbalance between the two (Logothetis and Wandell, 2004; Nair, 2005; Ekstrom, 2010; Kim and Ogawa, 2012); see Figure 2.6. In other words, the amount of

deoxygenated haemoglobin shows initially a rapid increase (peaking at about 2 s after stimulus onset) as a result of oxygen consumption by active neurons, followed by a superfluous perfusion of oxygenated blood that exceeds metabolic need (peaking at about 6 s from stimulus onset). This flushes deoxy-haemoglobin from the venous system. Further details of the neurovascular coupling are not yet fully understood. For example, the complex neuronal versus astrocyte signalling mechanisms, the contributions of excitation versus inhibition, and the spatial extent and dynamic properties of neural versus vascular responses (e.g., Harrison et al., 2002; Uğurbil et al., 2003; Sheth et al., 2005; Kim et al., 2010; Handwerker et al., 2012; Fukuda et al., 2016; Goense et al., 2016; Winder et al., 2017; Uludağa and Blinder, 2018). Furthermore, Takata et al. (2018) recently reported that astrocytic activation without accompanying neuronal activation is able to evoke the fMRI BOLD signal, which strongly suggests that caution should be exercised when appraising the findings from fMRI studies.

2.7 Biophysical foundations of TMS

Transcranial magnetic stimulation (TMS) is a powerful technique that allows us to draw causal brain-behaviour inferences by interfering with the neural activity in specific cortical regions through the delivery of magnetic pulses (Pascual-Leone et al., 1999, 2000; Walsh and Rushworth, 1999; Walsh and Cowey, 2000; Devlin and Watkins, 2007). It operates on Faraday's law of electromagnetic induction (Barker et al., 1985). The TMS stimulator releases an electric current that circulates at high speed in the stimulation coil's wire loops. This gives rise to a time changing magnetic field which passes through the subject's scalp and skull with minimal attenuation, inducing an electric field in the underlying neural tissue that modulates cortical excitability (Wagner et al., 2009; Peterchev et al., 2012); see Figure 2.7A. Importantly, the direction of the induced current is perpendicular to the coil surface, with the locus of stimulation depending on the stimulating coil geometry and placement (Wagner et al., 2009; Peterchev et al., 2012); see Figure 2.7B. A trade-off between the focality and depth of the TMS-induced electric field is typically generated with a figure-of-eight coil configuration. This is the preferred

choice among researches because of its good spatial specificity at the intersection of the two loops (Deng et al., 2013). Irrespective of coil geometry (see Figure 2.7C), TMS is usually not well suited for targeting deeper brain structures.

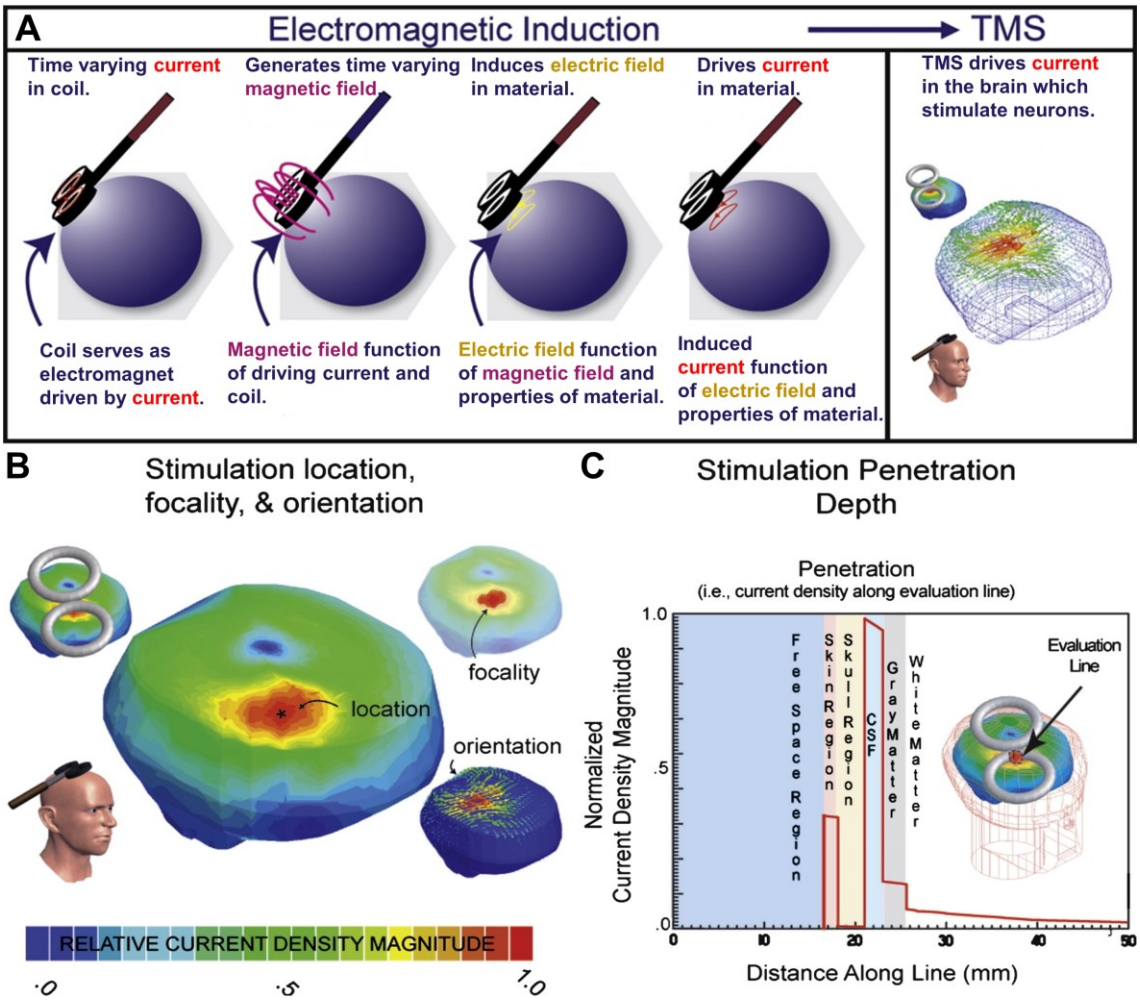

Figure 2.7: Biophysical foundations of TMS. (A) The electromagnetic principles that underlie TMS. (B) The location of stimulation, the area of stimulation and the current density orientations are determined by factors such as coil geometry and placement. (C) The current density magnitude varies with scalp-to-target distance, indicating that the degree of penetration of TMS pulses is limited. Although increasing magnetic pulse intensity improves stimulation depth, this is at the expense of stimulation focality. Adapted with permission (Elsevier) from Wagner et al. (2009).

When applied as repetitive trains of magnetic pulses at high frequency (10-20 Hz) as opposed to single pulses (i.e. repetitive transcranial magnetic stimulation), TMS introduces noise into local neural processing thereby causing a transient “virtual lesion” that typically lasts from milliseconds to minutes (Rossi et al., 2009; Miniussi et al., 2013; Valero-Cabré et al., 2017). Variants of the same TMS protocol involve the application of short bursts of 50 Hz rTMS in the theta range (i.e. 5 Hz), also known as patterned

repetitive TMS or theta-burst stimulation (TBS), which is becoming increasingly popular due to its ability to exert long-lasting disruptive effects (Rossi et al., 2009; Valero-Cabré et al., 2017); see Figure 2.8. Crucially, the combination of TMS with online stereotaxic neuronavigation systems allow specific cortical regions of the cortex to be perturbed by tracking the position of the TMS coil on a 3D reconstruction of each individual participant's structural MRI brain volume (Valero-Cabré et al., 2017). From a neurophysiological point of view, on the other hand, it is likely that there is no single mechanism that can explain the disruptive effect of TMS (Polanía et al., 2018). However, it is commonly agreed that TMS-induced "virtual lesions" may involve long-term-depression-like plasticity effects at the level of NMDA receptors (Dayan et al., 2013).

The difficulty in establishing a unitary neurophysiological mechanism of action of TMS is due to the complex interaction of a range of experimental factors including stimulation frequency, intensity, timing and duration, in addition to pulse train length, coil orientation, coil type, scalp-to-cortex distance and brain state at the time of stimulation (Pell et al., 2011; Sandrini et al., 2011; Peterchev et al., 2012; Parkin et al., 2015; Rossini et al., 2015; Valero-Cabré et al., 2017).

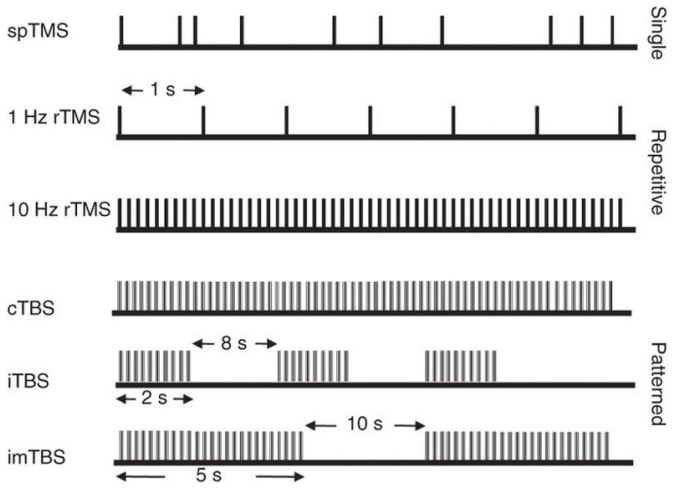

Figure 2.8: TMS protocols. TMS can be applied as single (spTMS) or multiple pulses in low (1 Hz rTMS) or high frequencies (10 Hz rTMS). More recent protocols such as theta-burst stimulation (TBS) are increasingly used, both in research and for clinical applications. TBS entails applying short bursts of 50 Hz pulses in the theta range as a continuous (cTBS; 20-40 s), intermittent (iTBS; 2 s interleaved by a pause of 8 s), or intermediate (imTBS; 5 s interleaved by a pause of 10 s) train. Contrary to cTBS, iTBS is associated with facilitatory after-effects. Adapted with permission (Springer Nature) from Dayan et al. (2013).

2.8 Prediction accuracy

In line with the primary goal of the current thesis, the experimental chapters assess the predictive power of a set of regions of interest informed by prior fMRI and TMS studies of neurologically-normal subjects. By arranging all patients in a 2×2 contingency table according to the presence or absence of damage to a region of interest and the presence or absence of a functional impairment of interest (for details, see experimental chapters), prediction/classification accuracy can be expressed in terms of positive and negative predictive values, sensitivity, specificity and the odds ratio (Altman and Bland, 1994a, b; Bland and Altman, 2000; Glas et al., 2003). Critically, these measures allow the strength of the relationship between a binary predictor/classifier (i.e. the presence or absence of damage) and the behaviour it is trying to predict/classify (i.e. the presence or absence of a functional impairment) to be evaluated from a multitude of angles. As can be seen in Table 2.1, each of the accuracy metrics mentioned above are derived as follows:

Table 2.1: 2×2 contingency table.

		Functional Impairment	
		Present	Absent
Region of Interest	Damaged	TP	FP
	Preserved	FN	TN

TP = true positives; FP = false positives; FN = false negatives; TN = true negatives.

(i) **Positive Predictive Value** = TP / (TP + FP)

The positive predictive value indicates the proportion of patients with the functional impairment of interest among those with damage to the region of interest.

(ii) **Negative Predictive Value** = TN / (FN + TN)

The negative predictive value indicates the proportion of patient without the functional impairment of interest among those with no damage to the region of interest

(iii) **Sensitivity** = TP / (TP + FN)

The sensitivity indicates the proportion of patients with damage to the region of interest among those with the functional impairment of interest.

(iv) **Specificity** = TN / (FP + TN)

The specificity indicates the proportion of patients with no damage to the region of interest among those without the functional impairment of interest.

(v) **Odds Ratio** = (TP ÷ FP) / (FN ÷ TN)

The odds ratio indicates how many times higher the odds for the presence of the functional impairment of interest is in those with damage to the region of interest than those with no damage to the region of interest. In other words, it represents a single metric that reflects the overall performance of a binary predictor/classifier (i.e. its calculation considers all four cells of the 2×2 contingency table). In contrast to other single indicators of predictor/classifier performance such as overall accuracy or Youden's index, the odds ratio is prevalence-independent and has a more straightforward interpretation (Glas et al., 2003). In addition, it is widely used in the biomedical literature, which makes it preferable to potential alternatives because the current work aims to reach biomedical researchers.

CHAPTER 3 (Experiment I):

Predicting Phonological Processing Impairments from Stroke Damage to Regions Identified by TMS and fMRI Studies of Normal Phonological Processing

3.1 Summary

Transcranial magnetic stimulation (TMS) focused on either the left supramarginal gyrus (SMG) or opercular part of the left inferior frontal gyrus (pOp) has been reported to transiently disrupt the ability to perform phonological more than semantic tasks in neurologically-normal individuals. Likewise, many functional magnetic resonance imaging (fMRI) studies have shown greater activation in SMG and/or pOp for phonological relative to semantic tasks. Here I sought to determine whether damage to the regions identified by previous TMS and fMRI studies impaired phonological processing abilities in right-handed, English speaking adults, who were investigated at least 1 year after a left-hemisphere stroke.

When the regions of interest were limited to 0.5 cm^3 of grey matter centred around sites that had been identified with TMS-based functional localisation, phonological impairments were observed in 74% (40/54) of patients with damage to the SMG and/or pOp regions and 21% (21/100) of patients sparing these regions. This classification accuracy was significantly better than that observed when using regions of interest centred on activation sites from prior fMRI studies of normal phonological processing (odds ratio = 10.7 for TMS versus 7.0 for fMRI). Moreover, the discriminatory power of the functionally localised TMS sites was not improved further by combining TMS and fMRI sites (odds ratio = 9.5). These results suggest that TMS-based functional localisation in neurologically-normal participants helped to identify which parts of pOp and SMG were necessary for efficient phonological processing. More generally, the use of regions of interest derived from TMS studies of neurologically-normal individuals might, in the future, help us to improve our ability to predict outcome and recovery after stroke.

3.2 Introduction

The goal of my first study was to investigate whether, and how consistently, phonological processing abilities are impaired by stroke damage to regions associated with normal phonological processing by prior TMS and fMRI studies. There are several reasons why damage to regions derived from TMS, fMRI and lesion studies might have inconsistent effects on phonological processing abilities. For example, fMRI and TMS studies of neurologically-normal subjects typically establish a region's contribution to task performance based on inferred neuronal activation (in fMRI studies) or slowed reaction times during disruptive stimulation (in TMS studies). However, neither of these measurements indicate that the region is essential for accuracy. They may not be if other intact brain areas can take over, or learn to take over, the lost function. Lesion studies are therefore needed to assess how critical a region is for a given type of processing. Similarities between TMS-induced "virtual" lesions and stroke-induced "real" lesions are, on the other hand, that minor forms of damage may lead to slower reaction times rather than reduced accuracy; and in both cases there can be neuroplasticity-related changes at the network level (Hartwigsen, 2018). For example, Price et al. (1999) reported a patient who, despite not showing an impairment of accuracy in the context of stroke damage to a region of interest (ROI), was, nonetheless, significantly slower than normal. In contrast, the challenge with lesion studies is that they might not necessarily locate the regions where damage is causing the deficit, particularly when (i) the cohort studied includes patients with large lesions and (ii) the inferred critical locus is biased towards areas that are most susceptible to vascular events (Kimberg et al., 2007; Inoue et al., 2014; Mah et al., 2014; Sperber and Karnath, 2017). The motivation for my first study was thus to determine whether TMS and fMRI could be used to help localise the brain regions where stroke damage consistently and persistently impairs accuracy on phonologically demanding tasks.

Previous research in neurologically-normal individuals has shown that targeting either the left supramarginal gyrus (SMG) (Hartwigsen et al., 2010a; Sliwinska et al.,

2015) or pars opercularis of the left inferior frontal gyrus (pOp) (Gough et al., 2005; Hartwigsen et al., 2010b) with repetitive transcranial magnetic stimulation (TMS) increases reaction times during tasks taxing phonological processing (e.g., do two words sound the same?) more than during tasks indexing semantic processing (e.g., are two words related in meaning?). These TMS findings are consistent with a wealth of functional imaging studies that have shown greater activation in SMG and/or pOp for phonological relative to semantic tasks (Price et al., 1997; Poldrack et al., 1999; Booth et al., 2002; Devlin et al., 2003; McDermott et al., 2003; Seghier et al., 2004; Gitelman et al., 2005; Simard et al., 2013).

In the current study, it was hypothesised that, if SMG and pOp are necessary for normal phonological processing, then stroke damage to these regions should impair the ability to perform phonologically demanding tasks, unless other brain regions (e.g., in the right hemisphere) can compensate for the lost function. More generally, if this proof-of-principle study finds that damage to the same regions associated with phonological processing in prior TMS and/or fMRI investigations also impairs phonological processing in stroke patients, then TMS and/or fMRI may provide a useful tool for identifying and understanding lesion sites that predict outcome after stroke.

In this context, it is relevant to make a distinction between regions of interest that were previously identified using TMS-based functional localisation (Gough et al., 2005; Sliwinska et al., 2015) and TMS sites that were centred on areas of peak activation from previous fMRI studies (Hartwigsen et al., 2010a, b). This is because TMS-based functional localisation has shown that the site that is most sensitive to TMS (averaged over subjects) does not necessarily correspond to the site of peak activation in fMRI studies (averaged over subjects). For example, the average TMS pOp site [-52, 16, 8] identified after functional localisation in Gough et al. (2005) is at least 1 cm anterior and inferior to the fMRI coordinates for pOp activation [-50, 6, 24] reported in Devlin et al. (2003), with an Euclidean distance of 19 mm between the two points. Likewise, the average TMS SMG site [-52, -34, 30] identified after functional localisation by Sliwinska

et al. (2015) is at least 1 cm inferior and lateral to the fMRI coordinates for SMG activation [-42, -40, 46] reported in Devlin et al. (2003), with an Euclidean distance of 20 mm between the two points. Conversely, the TMS sites reported by Hartwigsen et al. (2010a, b) without functional localisation ([-47, 6, 21] for pOp and [-45, -39, 45] for SMG) are almost identical to those reported in the fMRI study by Devlin et al. (2003).

Previous lesion studies have already shown that stroke damage to SMG and/or pOp and/or the white matter underlying these areas can cause deficits in tasks that require phonological processing, including speech production (Mirman et al., 2015), sentence production (Faroqi-Shah et al., 2014), non-word repetition (Faroqi-Shah et al., 2014), and phonological decisions (Geva et al., 2011). Additionally, the number of phonological errors produced during picture naming has also been reported to correlate with lesion status in a left perisylvian region involving SMG, pOp and the withe matter underneath (Schwartz et al., 2012). However, all the lesion studies mentioned above averaged effects across groups of patients including those with large lesions that damaged multiple brain regions. It is therefore not possible to establish whether phonological processing was impaired by damage to SMG alone, pOp alone or a combination of both. Nor does it allow us to determine how consistently damage to either of these regions leads to phonological processing impairments in individual subjects. Moreover, other studies that have dissociated the brain areas where tissue loss is uniquely related to phonological or semantic processing impairments (Butler et al., 2014; Halai et al., 2017) identified SMG and pOp sites that were not exactly the same as those targeted in the TMS studies that were used to inform the current experiment (i.e. Gough et al., 2005 and Sliwinska et al., 2015). This inter-study variability in the lesion sites associated with phonological processing impairments might be the consequence of (i) inter-patient differences in the effect of damage to the same areas (Lorca-Puls et al., 2018); (ii) inter-study differences in the type of phonological processing being tested; (iii) inter-study differences in the lesion sites being tested (Lorca-Puls et al., 2018); and/or, (iv) the possibility that damage to multiple brain regions can cause the same impairment

with only a subset of them detected in any given lesion analysis (Gajardo-Vidal, Lorca-Puls et al., 2018).

To better understand inter-patient variability, lesion studies need to (i) determine when and where stroke damage consistently impairs phonological processing abilities following SMG and pOp lesions; (ii) evaluate the effect of damage to different parts of SMG independently from damage to different parts of pOp (and vice versa); and (iii) compare how well the behavioural effect of interest is captured by a variety of phonological processing tasks.

In summary, Experiment I aimed to investigate the consistency of lesion effects in individual patients. In particular, how well the presence or absence of damage to a set of regions of interest predicts the presence or absence of a deficit of interest (i.e. classification/prediction accuracy). This is essential for testing our ability to generate accurate outcome predictions for future patients. In contrast, many lesion studies focus on the statistical significance of group-level effects (i.e. averaged over subjects), rather than the consistency of effects across individual patients. The problem with group-level statistics is that (a) they treat what might otherwise be meaningful between-subject variability as noise in the data (Seghier and Price, 2018) and (b) statistical significance in lesion analyses can be driven by inconsistent effects that are only present in a subset of patients (Price et al., 2017; Gajardo-Vidal, Lorca-Puls et al., 2018; Lorca-Puls et al., 2018). By focusing on the consistency of lesion effects rather than their statistical significance, it is therefore expected that such issues will be (at least partially) avoided.

3.2.1 Research question

In brief, the current experiment attempted to address the following research questions:

- Can damage to regions from prior TMS and fMRI studies of phonological processing in neurologically-normal individuals be used to accurately predict the incidence of persistent phonological impairments in stroke patients?

- If so, which of the two types of regions of interest (from prior TMS or fMRI studies) yields the best classification accuracy?

3.3 Materials and Methods

3.3.1 Participants

Data from all participants were retrieved from the Predicting Language Outcome and Recovery After Stroke (PLORAS) database (Price et al., 2010a; Seghier et al., 2016); for details, see Chapter 2. At a minimum, the data available for each patient included: a full assessment of speech and language abilities using the Comprehensive Aphasia Test (Swinburn et al., 2004); and a 3D lesion image, in standard space, created from a T1-weighted high resolution (1 mm isotropic voxels) anatomical whole-brain volume, using an automated lesion identification software (Seghier et al., 2008a). The study was approved by the Joint Research Ethics Committee of the National Hospital for Neurology and Neurosurgery and the UCL Queen Square Institute of Neurology. All patients gave written informed consent prior to participation and were compensated £10 per hour for their time.

Two samples of participants were selected: Sample 1 was used to investigate phonological processing abilities after damage to regions of interest from TMS and fMRI studies; Sample 2 was used to select phonologically demanding tasks that had been administered to all the patients in Sample 1. The following inclusion criteria were common to both samples:

(1) aged over 18;

(2) no history of neurological or psychiatric illness (other than stroke);

(3) right-handed (pre-morbidly), to minimise confounds from an increased probability of atypical language lateralisation in left-handers (e.g., Knecht et al., 2000; Szaflarski et al., 2002);

(4) normal or corrected-to-normal vision and hearing (as per self-reports), to ensure that impaired performance on any of the language tasks was not a consequence of the inability to see and/or hear the stimuli;

(5) native speaker of English (either monolingual or bilingual), to ensure that lower scores were not the consequence of pre-morbid language abilities and to control for differences in language recovery trajectories between monolingual native English speakers and bilingual non-native English speakers following stroke (as shown in Hope et al., 2015);

(6) at least 1 year post-stroke, to allow for changes in cerebral autoregulation and functional reorganization to have largely occurred (e.g., Saur et al., 2006); and

(7) more than 1 cm^3 of stroke damage (as measured by an automated lesion identification tool), because (i) the binary 3D lesion images used in the current study (for details, see Chapter 2) were not available for patients whose lesions were smaller than 1 cm^3 and (ii) lesions smaller than 1 cm^3 seldom result in phonological impairments (Price et al., in preparation); therefore, their inclusion would inflate the true negative rate (i.e. no deficit when little damage) for both types of prediction (from TMS and fMRI regions of interest).

In addition, the following inclusion criteria were specific to Sample 1:

(a) left-hemisphere stroke (as attested by a clinical neurologist), to focus on patients whose language outcomes are unknown as opposed to patients with right-hemisphere lesions who were not expected to have phonological impairments; and

(b) up to 5 years post-stroke, to control for longer term changes related to slow recovery or decline caused by vascular dementia (the incidence of which has been shown to be higher in stroke patients than neurologically-normal individuals; Pendlebury and Rothwell, 2009; Allan et al., 2011; Gorelick et al., 2011; O'Brien and Thomas, 2015; van der Flier et al., 2018).

A total of 154 patients met these criteria and were included in Sample 1 (see Table 3.1 and Figure 3.1). The neuroimaging and behavioural data from these patients were used to test how consistently damage to very specific regions of interest (from TMS and fMRI) resulted in phonological processing impairments.

Table 3.1: Summary of demographic and clinical data for Sample 1.

Factor		Sample 1 (*N* = 154)
Age at stroke onset (years)	*M*	56.2
	SD	12.7
	Range	18.7-86.5
Age at scan acquisition (years)	*M*	59.0
	SD	12.7
	Range	21.3-90.0
Time post-stroke (years)	*M*	2.7
	SD	1.2
	Range	1.0-5.0
Lesion size (cm^3)	*M*	80.1
	SD	79.9
	Range	1.4-464.7
Gender	Males	111
	Females	43

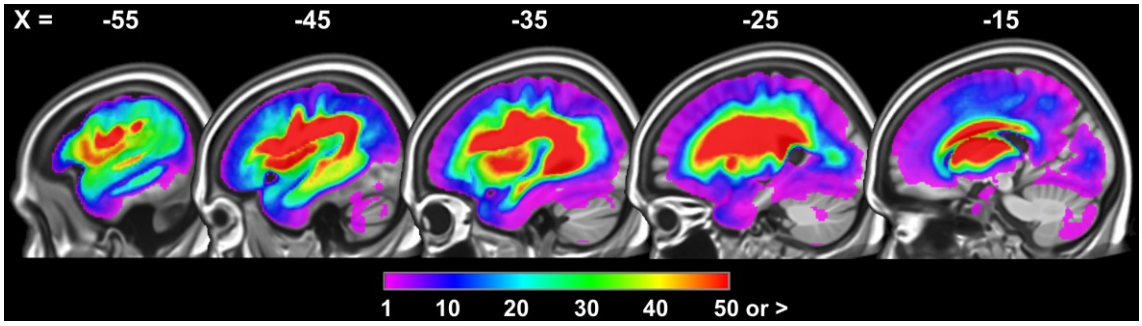

Figure 3.1: Lesion overlap map for Sample 1. Lesion overlap map of the 154 stroke patients included in Sample 1, depicting the frequency of overlapping lesions at each given voxel.

The second sample (i.e. Sample 2) was used to aid in the selection of phonological measures that showed a good sensitivity-specificity balance. To this end, Sample 2 was uniquely comprised of patients who had also been tested on the phonological and semantic decision tasks used in the TMS studies from which the regions of interest were derived (i.e. Gough et al., 2005 and Sliwinska et al., 2015); for details, see Chapter 2. By comparing their performance on the TMS tasks with that on the CAT tasks that had been administered to all patients (including those in Sample 1), it was possible to identify a set of phonological measures that best captured the variance in the TMS phonological task. A total of 42 patients were allocated to Sample 2 (see Table 3.2 and Figure 3.2). There was no overlap between Samples 1 and 2.

Table 3.2: Summary of demographic and clinical data for Sample 2.

Factor		Sample 2 (*N* = 42)
Age at stroke onset (years)	*M*	53.7
	SD	12.4
	Range	23.2-77.3
Age at testing (years)	*M*	63.4
	SD	12.2
	Range	28.0-84.0
Time post-stroke (years)	*M*	9.7
	SD	6.9
	Range	1.1-35.0
Lesion size (cm^3)	*M*	76.7
	SD	79.0
	Range	3.7-302.9
Gender	Males	27
	Females	15
Hemisphere damaged	Left	18
	Right	22
	Both	2
Delay between CAT and TMS tasks administration (months)	*M*	17.6
	SD	16.6
	Range	0.0-51.5

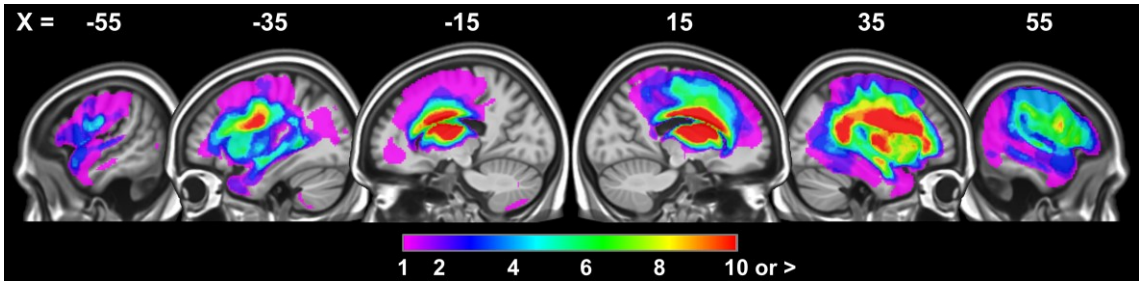

Figure 3.2: Lesion overlap map for Sample 2. Lesion overlap map of the 42 stroke patients included in Sample 2, depicting the frequency of overlapping lesions at each given voxel.

The patients in Sample 1 were, on average, significantly (i) younger at the time of data collection ($t(194) = 2.02$, $p = 0.045$) and (ii) earlier post-stroke ($t(42) = 6.58$, $p < 0.001$) than the patients in Sample 2. No other statistically significant between-sample differences in demographic and clinical variables were found. For formal comparison of CAT scores and lesion distributions, see Table 3.3 and Figure 3.3, respectively.

Table 3.3: Comparison of CAT scores between Sample 1 and Sample 2.

CAT Task No.	Sample 1 *M*(±*SD*)	Sample 2 *M*(±*SD*)	*t*	*df*	*p*
1	55.0(±9.0)	55.0(±9.4)	0.00	185	1.000
2	56.7(±5.8)	57.7(±4.7)	1.16	80	0.249
3	57.9(±10.7)	66.7(±6.3)	6.78	112	**0.000**
4	53.7(±7.4)	54.0(±5.5)	0.35	87	0.727
5	58.9(±7.9)	64.6(±5.1)	5.62	102	**0.000**
6	57.4(±7.4)	57.6(±7.5)	0.12	187	0.904
7	56.4(±6.9)	61.2(±4.0)	5.82	115	**0.000**
8	55.8(±7.7)	61.3(±4.7)	5.73	109	**0.000**
9	57.2(±8.7)	65.5(±5.8)	7.26	96	**0.000**
10	57.7(±9.1)	65.9(±4.7)	8.00	130	**0.000**
11	53.6(±8.1)	57.0(±5.7)	3.05	92	0.003
12	54.0(±8.8)	62.3(±5.0)	7.98	118	**0.000**
13	53.3(±10.0)	60.7(±3.6)	7.56	180	**0.000**
14	53.5(±8.8)	62.0(±6.1)	7.15	92	**0.000**
15	52.9(±9.1)	61.2(±6.1)	7.02	96	**0.000**
16	54.2(±9.4)	61.6(±3.6)	7.87	174	**0.000**
17	57.5(±11.0)	68.1(±5.3)	8.78	142	**0.000**
18	55.9(±10.2)	64.3(±6.4)	6.56	104	**0.000**
19	55.3(±9.5)	63.8(±6.3)	6.80	99	**0.000**
20	55.2(±9.6)	66.0(±4.9)	9.92	133	**0.000**
21	54.0(±11.4)	65.3(±5.3)	9.12	147	**0.000**
22	55.0(±9.8)	61.4(±2.8)	7.10	194	**0.000**
23	53.6(±10.7)	63.1(±7.2)	6.69	95	**0.000**
24	58.1(±6.0)	59.8(±4.0)	2.23	96	0.028
25	57.9(±9.0)	64.3(±4.6)	6.34	134	**0.000**
26	56.6(±9.1)	63.4(±5.3)	6.16	115	**0.000**
27	60.5(±9.8)	69.3(±4.6)	8.13	147	**0.000**

The *p* values highlighted in bold (right-most column) survived a Bonferroni correction for multiple comparisons (i.e. $p < 0.002$). See Chapter 2 for which tasks correspond to each number. *t* = t statistic; *df* = degrees of freedom.

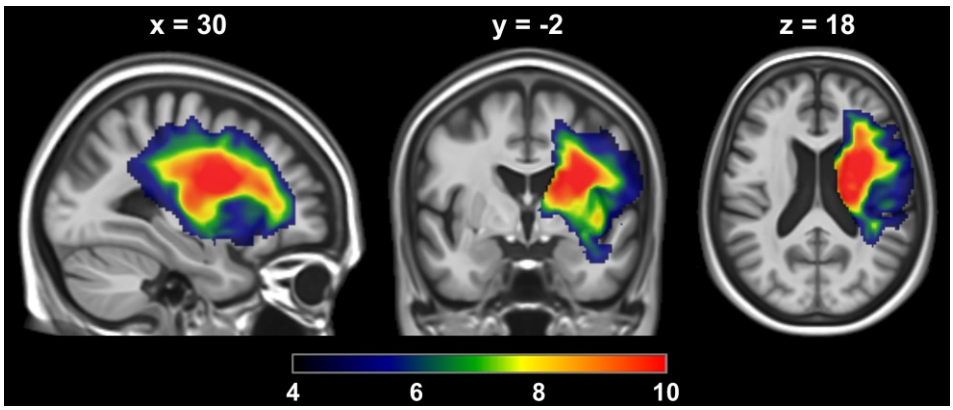

Figure 3.3: Comparison of the distribution of lesions between Sample 1 and Sample 2. The figure depicts areas of the brain where the degree of damage was significantly greater in patients from Sample 2 than Sample 1. As expected, the differences are restricted to the right-hemisphere given the composition of Sample 2 (i.e. 22 patients with right-hemisphere damage and 2 patients with bilateral damage) relative to that of Sample 1 (i.e. only patients with unilateral left-hemisphere damage). No brain regions were identified where the degree of damage was significantly greater in patients from Sample 1 than Sample 2. The statistical map comprises voxels that survived a voxel-level family-wise error (FWE) corrected threshold (i.e. $p < 0.05$, FWE-corrected) as well as an extent threshold for each cluster of 20 voxels. The colour scale shows the range of *t* values associated with statistically significant effects.

3.3.2 Assessing phonological processing abilities

All patients recruited to the PLORAS database are assessed on the Comprehensive Aphasia Test (CAT) (Swinburn et al., 2004). The CAT is a fully standardised test battery, which consists of a total of 27 different tasks; for details, see Chapter 2.

The current study focused on 3 tasks from the CAT that (i) required phonological processing while (ii) placing minimum demands on semantics. These tasks were “non-word reading”, “non-word repetition” and “digit span”. Patients with phonological processing impairments were expected to have difficulties performing all three of these tasks. In contrast, if performance was only impaired on the non-word reading task, this could be the consequence of poor visual perceptual/orthographic processing abilities. Likewise, if performance was only impaired on the non-word repetition and/or digit span tasks, this could be the consequence of poor auditory perceptual/speech processing skills.

There are no “pure” indices of phonological and semantic processing. For example, the CAT non-word reading, non-word repetition and digit span tasks arguably involve multiple cognitive component processes such as translating visual inputs (for non-word reading) or auditory inputs (for non-word repetition) into phonological representations, and holding these phonological codes in working memory prior to speech articulation (for digit span). Similarly, the TMS phonological and semantic tasks can be broken down into a number of levels of processing, each drawing on specific computations and representations. For instance, the TMS semantic task requires comparing the meaning of two written words online, which may in part be supported by implicitly accessing the associated phonological representations (although to a much lesser extent than the role played by the respective semantic representations). In consideration of this situation, I performed a detailed hypothesis-driven task analysis to inform the selection of tasks that differentially weight phonological versus semantic processing abilities (see Figure 3.4). In addition, in order to validate the tasks utilised to ascertain the presence or absence of phonological processing impairments in the current lesion study, a direct comparison of CAT and TMS phonological scores was conducted (see Section 3.3.3). These issues will be revisited in Chapter 6.

3.3.3 Cross-validating the definition of phonological processing impairments

The 3 phonologically demanding tasks selected from the CAT are different to the tasks used to define the regions in the TMS studies by Gough et al. (2005) and Sliwinska et al. (2015); for a task analysis, see Figure 3.4. To ensure that comparable types of phonological processing were being tested, I used data from Sample 2 to correlate scores on the TMS phonological and semantic tasks (from Gough et al. 2005 and Sliwinska et al. 2015; see Section 3.3.4 for why these studies were selected to inform the ROI definition step) with scores on each of the three phonologically demanding tasks from the CAT (i.e. non-word reading / non-word repetition / digit span), and every possible combination of them.

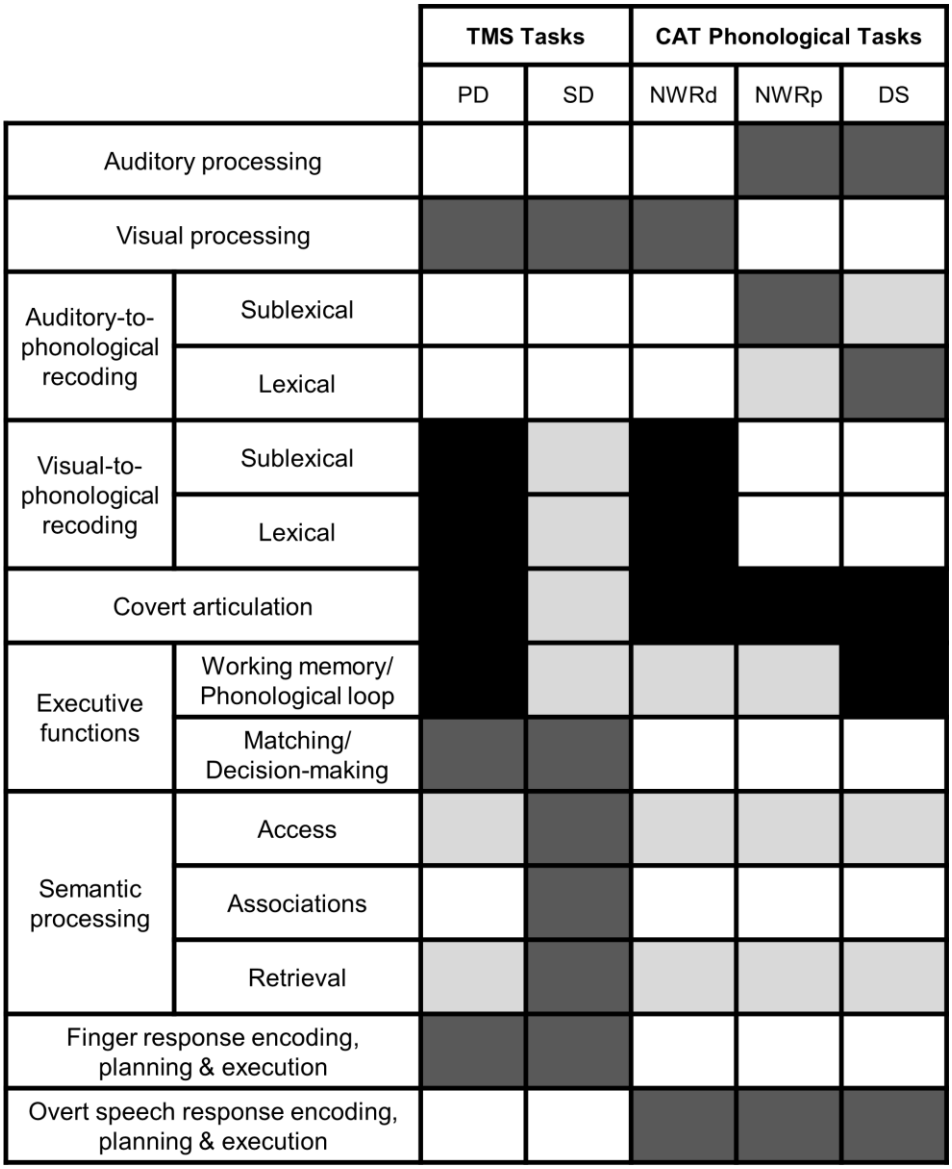

Figure 3.4: Task analysis. The levels of processing hypothesised to be required for completing the TMS phonological (PD) and semantic (SD) decision tasks, and the following tasks from the Comprehensive Aphasia Test (CAT): non-word reading (NWRd), non-word repetition (NWRp) and digit span (DS). Black is used to highlight the phonological processes of interest that are shared by the TMS phonological task and at least one of the CAT phonological tasks. Dark grey indicates necessary/explicit processes. Light grey signifies supporting/implicit processes.

A significant correlation between the TMS and CAT phonological scores would suggest that both tapped phonological processes that are shared by a range of language tasks. This is preferable to two potentially attractive alternatives. One would be to conduct a new TMS study in neurologically-normal participants using the CAT tasks. The problem here is that all 3 of the CAT tasks that selectively weight phonological processing involved speech production, which is difficult to measure in TMS studies because voice onset times are hidden by noise and jaw movements. The other option

would be to administer the TMS phonological task to the patients in Sample 1. Here the problem would be that a significant number of individuals from Sample 1 would not be able to perform the task or generate enough correct responses that can be used to measure reaction times.

3.3.4 Regions of interest from prior TMS and fMRI studies

Only those TMS studies that met the following inclusion criteria were selected for ROI definition: (i) neurologically-normal right-handed participants were tested in their native language; (ii) TMS was delivered over the left supramarginal gyrus (SMG) and/or the pars opercularis of the left inferior frontal gyrus (pOp); (iii) TMS-based functional localisation was used to identify the SMG and/or pOp sites that are most sensitive to disruptive stimulation prior to the main experiment; (iv) performance on well-matched phonologically and semantically demanding tasks was directly compared; (v) TMS over SMG and/or pOp had a significantly greater disruptive effect on phonological than semantic processing abilities; (vi) the average coordinates of the SMG and/or pOp testing sites where TMS resulted in statistically significant effects were provided in MNI or Talairach space. The decision to exclusively focus on studies that utilised TMS-based functional localisation was motivated by the fact that, without functional localisation, stimulation sites in TMS studies are typically taken from previous fMRI studies (e.g., Hartwigsen et al. 2010a, b), which would render any comparison between the predictive power of TMS-derived and fMRI-derived regions redundant. Consequently, solely the TMS studies by Gough et al. (2005) and Sliwinska et al. (2015) were deemed suitable for the ROI definition stage.

In Gough et al. (2005) and Sliwinska et al. (2015), TMS-based functional localisation was achieved by delivering TMS over either pOp or SMG while neurologically-normal participants performed a phonological task in which they had to decide whether two written words rhymed (i.e. the localiser task). When TMS shortened response times relative to no TMS (i.e. had a facilitatory effect), another site within the same brain region was tested. When TMS increased response times (i.e. had a

disruptive effect), the same site was retested to determine whether stimulation resulted in consistent slowdowns. Only a site where TMS resulted in two or more response-time slowdowns during the localiser task was selected as a target site for that particular participant for the main experiment. In the main experiment, alternative phonological and semantic tasks were employed where subjects had to make judgements about the sound structure (e.g., do two words sound the same?) or meaning (e.g., are two words related in meaning?) of visually presented pairs of words.

In the current experiment, the TMS regions of interest were spheres (radius of 5 mm, 0.5 cm^3 in volume), centred on the mean MNI coordinates for SMG [-52, -34, 30] and pOp [-52, 16, 8] that were reported in studies that used TMS-based functional localisation (i.e. Gough et al., 2005 and Sliwinska et al., 2015); see Figure 3.5. The size of the regions (5 mm radius) was chosen based on the expected spatial resolution of TMS, which has been argued to be in the order of 5-10 mm (Brasil-Neto et al., 1992a, b; Wilson et al., 1993; Ravazzani et al., 1996; Thielscher and Kammer, 2002).

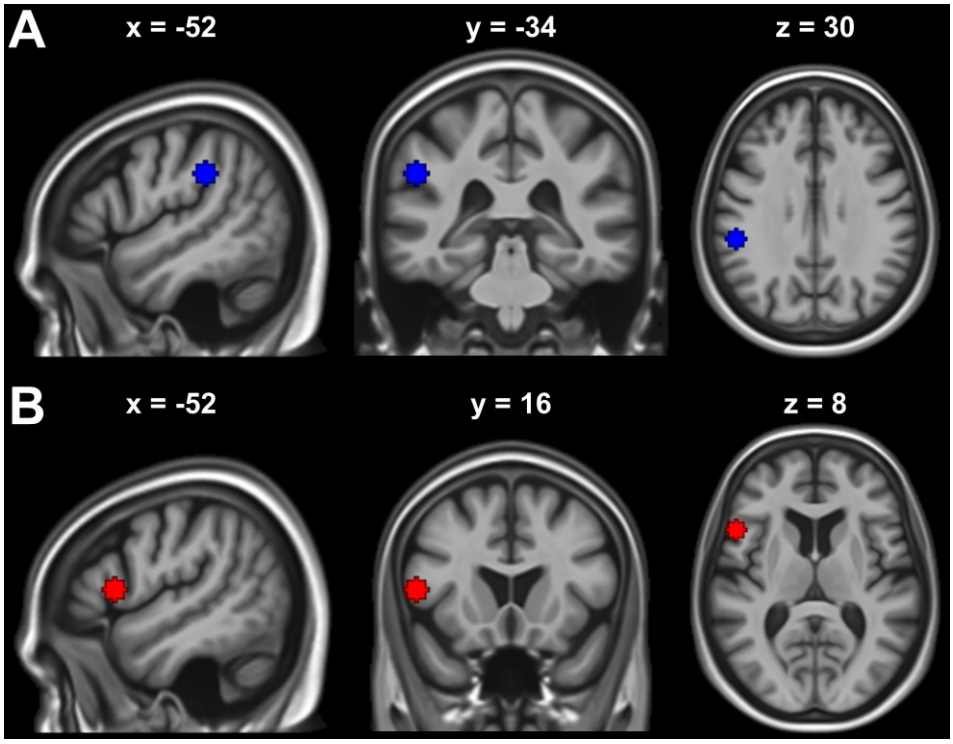

Figure 3.5: TMS regions of interest. The TMS regions are shown in blue for SMG (A) and red for pOp (B).

Regarding the inclusion criteria for fMRI study selection, the literature was searched for articles that: (i) employed an acquisition protocol which ensured whole-brain coverage; (ii) tested neurologically-normal right-handed participants in their native language; (iii) explicitly contrasted activation elicited by phonologically and semantically demanding tasks; (iv) reported statistically significant effects for the comparison of phonologically more than semantically demanding tasks in SMG and/or pOp; (v) provided the locations of peak activation within SMG/pOp in MNI or Talairach space. From those studies that met these criteria, only the coordinates of the most significant voxel within SMG and/or pOp (i.e. the one associated with the maximum statistic value) for the contrast "phonologically more than semantically demanding tasks" were used for ROI definition. It is worth noting that other researchers might prefer to conduct coordinate-based meta-analyses on fMRI data (e.g., Eickhoff et al., 2012) to derive potential critical regions. Here, I decided to focus on fMRI studies that directly investigated functional specialisation within SMG and pOp for phonological relative to semantic processing so as to maximise comparability between the selected fMRI and TMS studies in as many aspects as possible.

The six fMRI studies briefly described below were ultimately selected to inform the ROI definition step. In Booth et al. (2002), the phonological task required participants to determine whether a final word rhymed with either of two preceding words. For the semantic task, participants had to determine whether a final word was related in meaning with one of two preceding words. Different sets of words were presented in the visual and auditory modalities in separate runs. In Devlin et al. (2003), the phonological task required subjects to decide whether a written word had two syllables. For the semantic task, subjects had to decide whether a written word represented a man-made item. In Seghier et al. (2004), the phonologic task required participants to give a response only if two visually presented words rhymed. For the semantic task, participants had to respond whenever two visually presented words belonged to the same semantic category. In McDermott et al. (2003), subjects were instructed to attend closely to the relations among

upcoming written words. For the phonological task, they were told to think about how the words sounded alike. For the semantic task, they were told to think about how the words could be meaningfully connected (i.e. semantic associations). In Gitelman et al. (2005), the phonological task required participants to respond if two visually presented words were homophones. For the sematic task, participants had to respond if two visually presented words were synonyms. In Simard et al. (2013), subjects had to match a new word presented at the centre of the screen with one of four reference words presented at the top of the screen according to specific rules discovered by trial and error using feedback. The phonological rule required subjects to do the matching based on syllable rhyme, whereas the semantic rule required subjects to do the matching based on sematic category. Across all studies, the contrast of interest yielded brain regions that showed significantly greater activation for phonologically than semantically demanding tasks (irrespective of input modality).

In the current experiment, the fMRI regions of interest were also spheres (radius of 5 mm, 0.5 cm^3 in volume). For SMG, they were centred on x, y, z MNI coordinates [-57, -21, 39; -42, -40, 46; -54, -36, 40] from Booth et al. (2002), Devlin et al. (2003), and Seghier et al. (2004); see Figure 3.6. The SMG coordinates from Simard et al. (2013) were not considered because they fell outside the boundaries of the brain: [-60, -32, 52]. For pOp, the fMRI regions of interest were centred on [-58, 5, 13; -50, 6, 24; -57, 9, 24; -41, 3, 20] from McDermott et al. (2003), Devlin et al. (2003), Gitelman et al. (2005), and Simard et al. (2013); see Figure 3.7. The coordinates reported in McDermott et al. (2003) and Seghier et al. (2004) were converted from Talairach space to MNI space using the tal2icbm transformation (Lancaster et al., 2007).

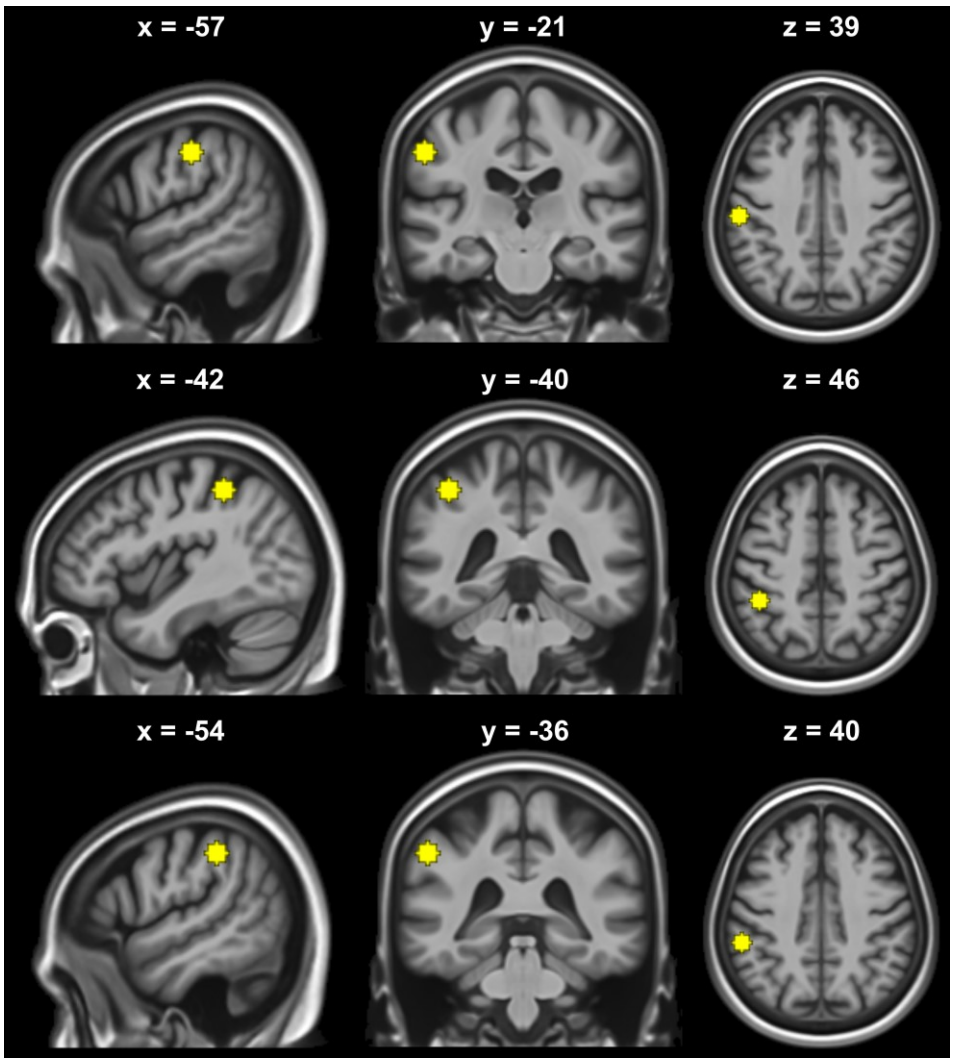

Figure 3.6: fMRI SMG regions of interest. From top to bottom, the fMRI SMG regions derived from Booth et al. (2002), Devlin et al. (2003), and Seghier et al. (2004) are shown in yellow.

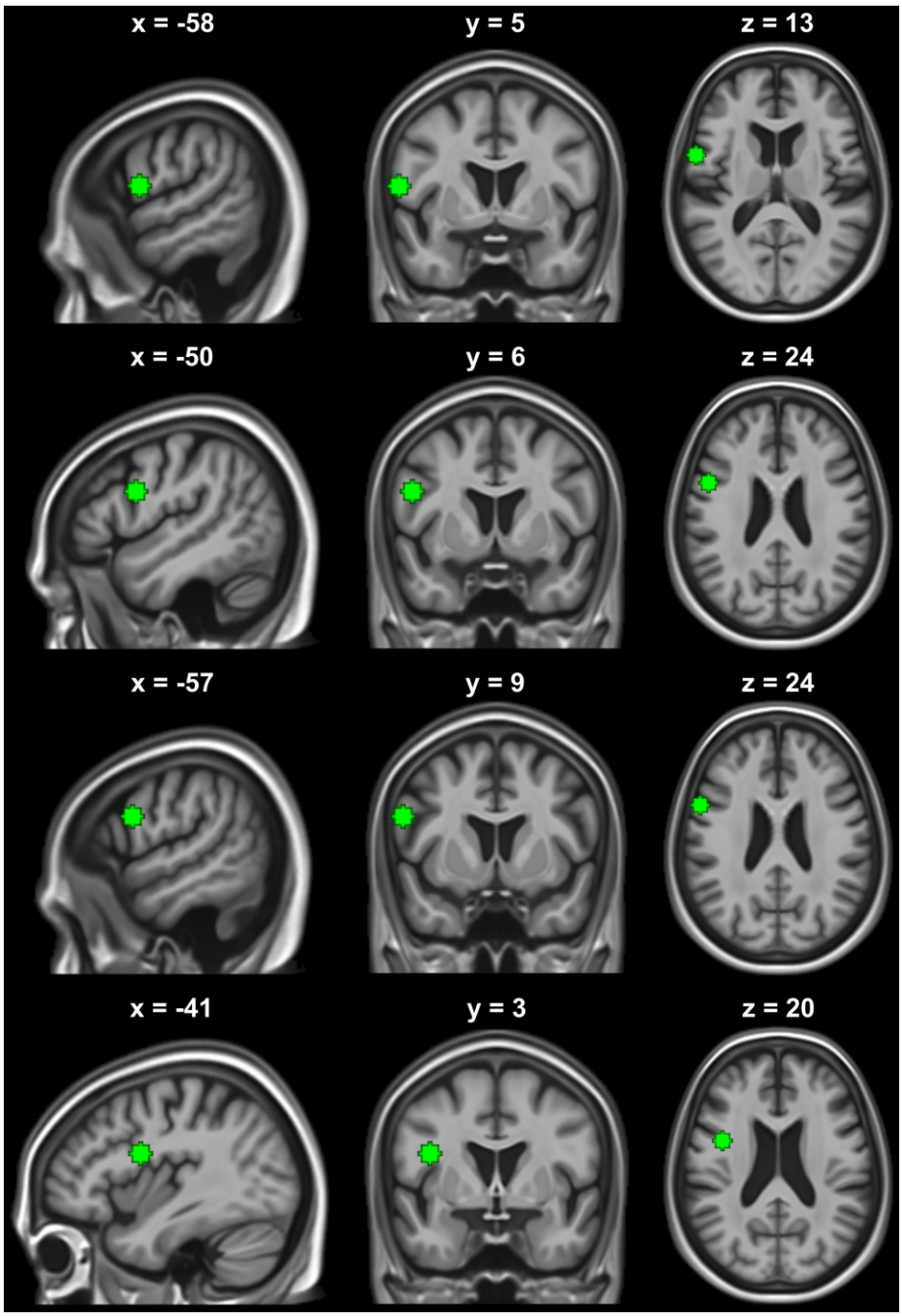

Figure 3.7: fMRI pOp regions of interest. From top to bottom, the fMRI pOp regions derived from McDermott et al. (2003), Devlin et al. (2003), Gitelman et al. (2005), and Simard et al. (2013) are shown in green.

3.3.5 Determining the threshold for critical damage

The experiment aims to establish whether damage to regions of interest from prior TMS and fMRI studies is consistently associated with phonological processing impairments in stroke patients. But, what constitutes damage? Does the whole of the region have to be damaged? If not, what proportion of damage consistently results in a

deficit? A section of the experiment therefore involved finding the percent damage "threshold" above which patients were more likely to have phonological processing impairments (as defined here) and below which patients were less likely to have phonological processing impairments. I established this threshold by finding the degree of damage that yielded the highest "odds ratio", which quantitatively describes the strength of the association between the presence/absence of one factor (e.g., phonological processing impairments) and the presence/absence of another factor (e.g., percent damage to a region of interest).

The first step in determining the threshold for critical damage was to divide all the patients in Sample 1 into those who did and did not present with selective phonological processing impairments (Subsets 1A and 1B). Selective phonological processing impairments were categorised as the occurrence of phonological impairments in the absence of semantic impairments. Semantics impairments were defined as impaired performance on a semantically demanding task that involved auditory inputs (i.e. auditory word-to-picture matching) in addition to a semantically demanding task that involved visual inputs (i.e. either visual word-to-picture matching or visual semantic associations). This subgrouping certainly depends on the sensitivity of the CAT tasks to detect phonological versus semantic processing impairments. With that being said, I chose to focus on patients with selective phonological processing impairments in an attempt to rule out the possibility that the critical damage thresholds were biased by those whose phonological processing impairments could be attributed to co-occurring auditory (i.e. impaired auditory word-to-picture matching) and visual perceptual deficits (i.e. impaired visual word-to-picture matching or visual semantic associations).

As the patients differed widely in lesion size, the second step in determining the threshold for critical damage involved matching the Subsets 1A and 1B for lesion size by finding the minimum and maximum lesion volumes that were common to both groups with no significant differences in mean lesion size across groups. The resulting group size was 23 patients in Subset 1A and 32 patients in Subset 1B (see Table 3.4 and

Figure 3.8). Of note, one potential risk of matching these subgroups for lesion size is that it could be the case that the likelihood of damage to any given critical region is very strongly correlated with the size of the lesion (i.e. the larger the lesion the more likely it is that the region of interest will be damaged), meaning that the statistical power of the analysis would be dramatically reduced. However, given the size of Sample 1 (N = 154), I considered this would be less of an issue than in smaller studies. Not matching the subgroups for lesion size (by range and mean) could, on the other hand, result in critical damage thresholds that are biased by patients with large lesions. In other words, critical damage thresholds that only indicate that the larger the lesion the more likely it is that the function of interest will be impaired, weakening the inferences that can be drawn with respect to the importance of lesion location. There is no perfect solution to the problem of teasing the effects of lesion size and site apart. However, in line with the aims of the thesis (i.e. predicting language outcome from lesion site), I decided to adopt a more conservative approach. In addition, an advantage of using a subset of the patients to determine the critical damage thresholds is that the remaining patients can serve as a validation group, which is somewhat similar to splitting the data into training and testing sets.

For each percentage of damage (i.e. 70%, 80%, 90% and 100%), the patients with and without phonological processing impairments were arranged into a 2×2 contingency table. This allowed me to express the classification accuracy in terms of positive and negative predictive values, sensitivity, specificity and the odds ratio (Altman and Bland, 1994a, b; Bland and Altman, 2000; Glas et al., 2003); for details, see Chapter 2. The comparison of predictive values for different regions was performed using the method proposed by Leisenring et al. (2000) as implemented in the R package “DTComPair” (Stock and Hielscher, 2014). In what follows, “damage” always refers to the proportion of damaged voxels among all voxels within specific regions of interest.

Table 3.4: Summary of demographic and clinical data for Subsets 1A and 1B.

Factor		Subset 1A (*n* = 23)	Subset 1B (*n* = 32)
Age at stroke onset (years)	*M*	51.1†	55.1†
	SD	14.9	11.7
	Range	18.7-75.4	24.8-73.4
Age at scan acquisition (years)	*M*	53.6†	58.1†
	SD	15.0	11.5
	Range	21.3-78.2	29.4-76.1
Time post-stroke (years)	*M*	2.4†	3.0†
	SD	1.0	1.1
	Range	1.0-4.5	1.1-4.7
Lesion size (cm^3)	*M*	82.2†	76.0†
	SD	25.4	25.5
	Range	44.3-128.7	44.0-135.7
Gender	Males	16	30
	Females	7	2

†No significant mean differences between Subset 1A and Subset 1B (i.e. $p > 0.05$).

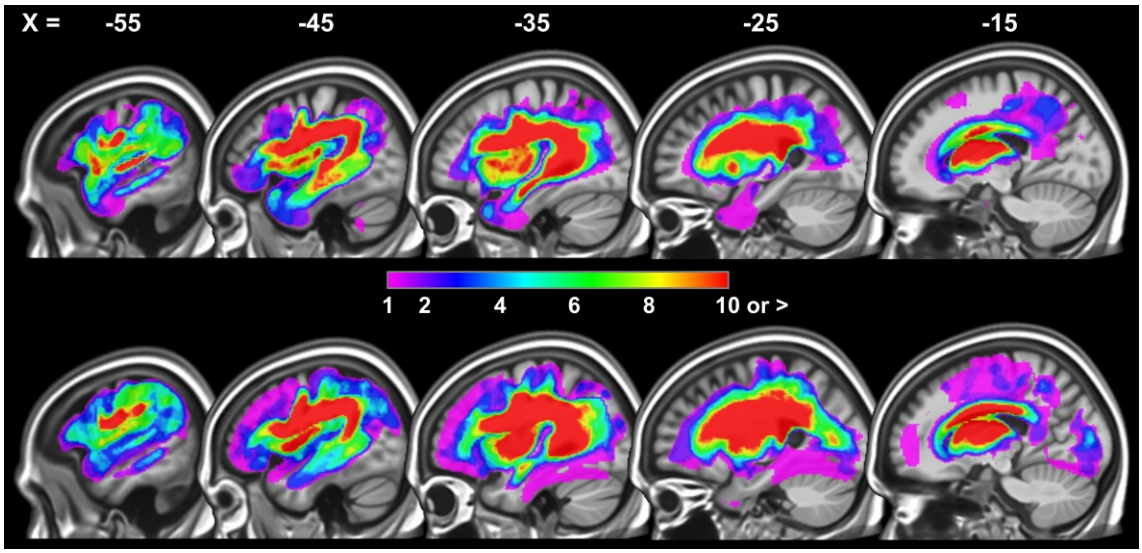

Figure 3.8: Lesion overlap map for Subsets 1A and 1B. Lesion overlap map of the 23 patients from Subset 1A (upper row) and the 32 patients from Subset 1B (bottom row), depicting the frequency of overlapping lesions at each given voxel. Comparison of the distribution of lesions between Subset 1A and Subset 1B did not yield any brain areas where the degree of damage was significantly greater in one subgroup relative to the other at a voxel-level statistical threshold of $p < 0.05$, FWE-corrected.

3.4 Results

3.4.1 Cross-validating the definition of phonological processing impairments

Figure 3.9 shows how strongly scores on the TMS phonological and semantic tasks correlated with scores on the three phonologically demanding tasks from the CAT (i.e. non-word reading / non-word repetition / digit span), and every possible combination of them, in the 42 patients from Sample 2 (i.e. those who had been tested on both the CAT and TMS tasks). The CAT measures that best captured the variance in the TMS phonological task were non-word reading ($r(40) = 0.64$) and the combination of non-word reading and digit span ($r(40) = 0.62$). Critically, both these phonological measures were also significantly more correlated with the TMS phonological task than the TMS semantic task ($z = 2.89$, $p = 0.004$ and $z = 3.00$, $p = 0.003$, respectively). In contrast, the association between the TMS phonological task and the CAT non-word repetition task did not reach statistical significance ($r(40) = 0.17$, $p = 0.288$). Therefore, "phonological processing impairments" were defined as the combination of abnormally low scores on both the non-word reading and digit span tasks.

The weak correlation between scores on the CAT non-word repetition task and the TMS phonological task (see Figure 3.9) may in part be explained by the fact that, in the CAT non-word repetition task, patients are provided with an auditory template of the target non-word which could be used as an external cue to guide speech articulation. This may lessen the demands on covert phonological processing prior to overt articulation, particularly when contrasted with the CAT digit span and non-word reading tasks.

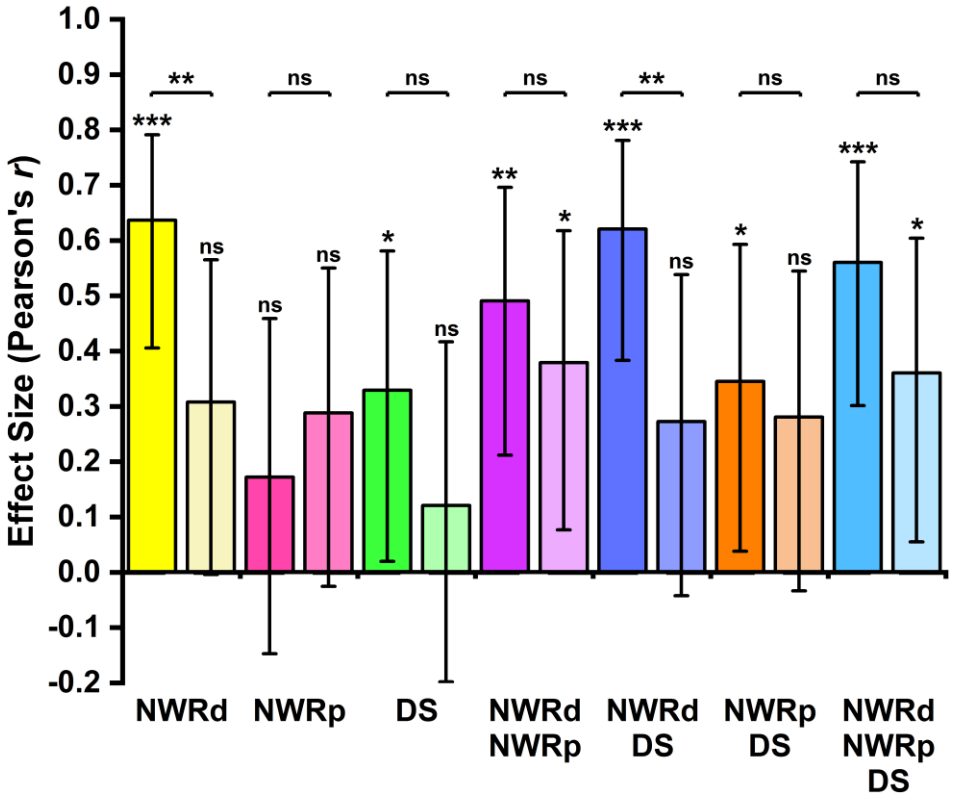

Figure 3.9: Strength of the association between TMS and CAT scores. The TMS phonological and semantic measures were the homophone judgement and semantic association tasks used by Gough et al. (2005) and Sliwinska et al. (2015). Scores on these tasks (for details, see Chapter 2) were correlated with those on seven different phonological measures from the CAT: non-word reading (NWRd), non-word repetition (NWRp), digit span (DS), and every combination of them (T-scores were averaged for each task pair/triplet). Two patients were classified as outliers because they had scores > 3 *SD* below the group mean on the non-word reading task and were therefore removed from all correlation analyses. For each CAT phonological measure, there are two columns: the left-hand side column indicates the strength of the correlation with the TMS phonological task, whereas the right-hand side column indicates the strength of the correlation with the TMS semantic task. Error bars represent 95% CIs. ns = not significant; * = statistically significant at $p < 0.05$; ** = statistically significant at $p < 0.01$; *** = statistically significant at $p < 0.001$.

3.4.2 Critical damage thresholds

The degree of damage that best explained the presence or absence of phonological impairments in Subsets 1A and 1B was 80%, for both the TMS and fMRI regions (see Table 3.5). Henceforth, “above-threshold damage” refers to 80% damage or above to any given set of regions.

Table 3.5: Critical damage thresholds for TMS and fMRI regions.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
TMS	100%	25%	57%	4%	91%	0.4
	90%	47%	60%	30%	75%	1.3
	80%	57%	68%	52%	72%	***2.8***
	70%	48%	63%	52%	59%	1.6
fMRI	100%	29%	54%	17%	69%	0.5
	90%	41%	57%	48%	50%	0.9
	80%	45%	64%	65%	44%	***1.5***
	70%	42%	58%	65%	34%	1.0

Threshold = percentage of the region of interest (ROI) damaged; PPV/NPV = positive and negative predictive values.

3.4.3 Classification accuracy for TMS regions

Summing over patients who had above-threshold damage to SMG, pOp or both, the incidence of phonological impairments was 74% (40/54). This was more than three times greater than the incidence of phonological impairments in patients who had below-threshold damage to both SMG and pOp (21/100 = 21%). When the classification accuracy of the TMS regions was re-expressed in terms of the odds ratio (i.e. a single metric that reflects the overall performance of a binary predictor/classifier), it was observed that the odds for the presence of phonological impairments in those with above-threshold damage to the TMS regions was 10.7 times higher than the odds for the presence of phonological impairments in those with below-threshold damage to the TMS regions (see Table 3.6).

When the effect of each lesion site was considered separately, I found that the incidence of phonological impairments was higher when SMG but not pOp was damaged (12/16 = 75%) than when pOp but not SMG was damaged (13/21 = 62%), and raised to 88% (15/17) when both SMG and pOp were damaged. In those with below-threshold damage to both regions, the incidence of phonological impairments was 14% (10/74) after 0-20% damage to either SMG or pOp, and 42% (11/26) after 21-79% damage to one or both of these regions.

Table 3.6: Classification accuracy for TMS and fMRI regions in Sample 1.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
(1) TMS	80%	74%	79%	66%	85%	10.7
(2) fMRI	80%	64%	80%	72%	73%	7.0
(1) & (2)	as above	64%	84%	80%	70%	9.5

Threshold = critical damage threshold for each region of interest (ROI); PPV/NPV = positive and negative predictive values.

3.4.4 Classification accuracy for fMRI regions

The TMS regions provided a better account of the data than the fMRI regions. For example, the odds ratio was 10.7 for the TMS regions compared to 7.0 for the fMRI regions. This indicates that the ability to discriminate between those with and without phonological impairments was higher when using the TMS regions than the fMRI regions. Moreover, combining the TMS and fMRI regions did not improve the classification accuracy relative to using the TMS regions only (9.5 versus 10.7); see Table 3.6.

The improved classification accuracy for TMS compared to fMRI predictions arose because the incidence of phonological impairments in the presence of above-threshold damage was significantly lower for the fMRI regions (44/69 = 64%) than the TMS regions (40/54 = 74%); X^2 = 4.51, p = 0.034. There were no differences in the incidence of phonological impairments in those who had below-threshold damage to the TMS and fMRI regions (21/100 = 21% for TMS and 17/85 = 20% for fMRI). See Table 3.7 for spatial overlap between TMS and fMRI regions.

Table 3.7: Overlap between spherical TMS and fMRI regions.

	TMS ROIs		fMRI ROIs						
	SMG	pOp	SMG			pOp			
	1	2	3	4	5	6	7	8	9
1	100%	0%	0%	0%	2%	0%	0%	0%	0%
2	0%	100%	0%	0%	0%	0%	0%	0%	0%
3	0%	0%	100%	0%	0%	0%	0%	0%	0%
4	0%	0%	0%	100%	0%	0%	0%	0%	0%
5	2%	0%	0%	0%	100%	0%	0%	0%	0%
6	0%	0%	0%	0%	0%	100%	0%	0%	0%
7	0%	0%	0%	0%	0%	0%	100%	31%	0%
8	0%	0%	0%	0%	0%	0%	31%	100%	0%
9	0%	0%	0%	0%	0%	0%	0%	0%	100%

The table shows the percentage of each of the regions listed at the top contained within each of the regions on the left-hand side. x, y, z MNI coordinates for regions of interest (ROIs): 1 = [-52, -34, 30]; 2 = [-52, 16, 8]; 3 = [-57, -21, 39]; 4 = [-42, -40, 46]; 5 = [-54, -36, 40]; 6 = [-58, 5, 13]; 7 = [-50, 6, 24]; 8 = [-57, 9, 24]; 9 = [-41, 3, 20].

3.4.5 The real extent of the lesions in patients with damage to the TMS regions and phonological impairments

Examination of the lesion sites revealed that all the patients with above-threshold damage to either of the TMS regions (n = 25) had lesions that were much larger than the 0.5 cm^3 grey matter spheres. For example, the smallest lesion in patients with phonological impairments after above-threshold damage to the TMS SMG region was 44.3 cm^3 and extended into the primary somatosensory cortex, posterior part of the superior temporal gyrus/sulcus, portions of the long gyrus of the insula and surrounding white matter (see Figure 3.10A below and Figure 4.4 in Chapter 4). Likewise, for the TMS pOp region, the smallest lesion associated with phonological impairments was 48.8 cm^3 and included the primary motor and premotor cortices, portions of the superior/middle temporal gyri, short gyri of the insula, anterior lentiform nucleus and surrounding white matter (see Figure 3.10B below and Figure 4.3 in Chapter 4).

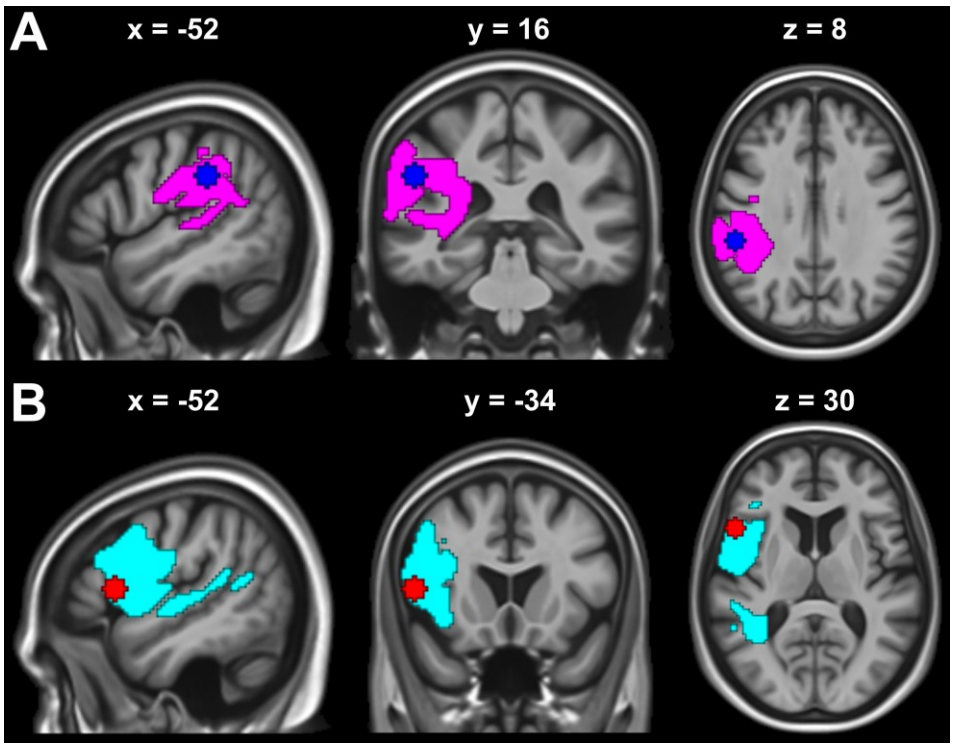

Figure 3.10: Smallest lesion sites. (A) The smallest lesion site associated with phonological impairments following above-threshold damage to the TMS SMG region (blue) is shown in violet (overlap = 100%). (B) The smallest lesion site associated with phonological impairments following above-threshold damage to the TMS pOp region (red) is shown in cyan (overlap = 87%). See also Figures 4.3 and 4.4 in Chapter 4 for lesion overlap maps.

3.5 Discussion

The current experiment studied the accuracy with which damage to TMS and fMRI regions of interest identified in neurologically-normal individuals predicted phonological processing abilities in patients with a left-hemisphere stroke. The regions of interest were 0.5 cm^3 spheres centred on MNI coordinates obtained from previous TMS and fMRI studies of normal phonological processing. The TMS sites were not the same as the fMRI sites (see Table 3.7) because they had been functionally localised to the position where disruptive TMS showed its maximal effect on phonological processing (Gough et al., 2005; Sliwinska et al., 2015). It was found that lesions to the TMS regions accounted for the incidence of persistent phonological impairments better than lesions to the fMRI regions and that the discriminatory power was not improved by combining the TMS and fMRI sites (see Table 3.6). This suggests that TMS-based functional

localisation in neurologically-normal participants helped to identify which parts of pOp and SMG were necessary for efficient phonological processing.

3.5.1 The predictive value of the TMS regions was surprisingly high

The high predictive value of the TMS regions is remarkable for the following three reasons. First, whereas TMS "virtual" lesion effects are transient (lasting typically from milliseconds to minutes), the effects of "real" lesions were measured at least 1 year after stroke onset which provides a substantial amount of time for recovery to occur. Second, phonological processing abilities were captured in the current lesion study using different tasks and less sensitive measures (accuracy versus response times) than those utilised in the original TMS studies (i.e. Gough et al., 2005 and Sliwinska et al., 2015). Third, the physiological mechanisms driving the effects of TMS and stroke on behaviour are known to differ. Specifically, TMS is thought to disrupt cortical excitability through long-term depression-like effects at the level of NMDA receptors (Dayan et al., 2013), whereas stroke (either ischemic or haemorrhagic) triggers a cascade of events that ultimately leads to neuronal cell death (Lo et al., 2003). In this context, the current study indicates that the effect of interest is sufficiently robust insofar as it was not completely masked by these potential confounding factors.

Additionally, the relatively high predictive power of the TMS regions implies that TMS may be a useful technique for guiding the identification of critical lesion sites in that it appears that the disruptive effect of TMS on behaviour tends to occur most frequently when regions that are necessary for task performance are stimulated (i.e. high specificity). It is therefore possible to imagine a situation where relatively weak forms of stimulation (or damage) lead to slowed reaction times, with the behavioural effects only flipping over into reduced accuracy when the stimulation (or damage) reaches a certain threshold. For example, Shimotake et al. (2015) investigated the role of the ventral anterior temporal lobe in semantic memory by delivering graded levels of stimulation intensity over a single functionally defined region in patients who had undergone subdural electrode implantation. The results revealed a significant positive relationship

between reaction times and stimulation intensity (i.e. slower reaction times as stimulation intensity increased) during picture naming and visual word-to-picture matching, with a decline in accuracy only being observed at the strongest intensity level.

3.5.2 The TMS regions are persistently necessary for accurate phonological processing

The effect of damage to TMS SMG or pOp on phonological processing was measured in terms of a persistent loss of accuracy on phonologically demanding tasks. In contrast, the TMS effects found in these regions were expressed in terms of a transient slowing of response times (Gough et al., 2005; Sliwinska et al., 2015). Slower response times as a result of disruptive stimulation in TMS studies indicate that the region of interest was involved in the process of interest but does not completely answer the question of whether the region was essential for task performance. Conversely, the fact that a loss of accuracy on phonologically demanding tasks was reliably observed after either TMS SMG or pOp damage suggests that other areas of the brain were not typically able to fully compensate for the function played by these regions, despite years of recovery time. In other words, the TMS regions seem to be persistently necessary for accurate phonological processing.

3.5.3 The lower predictive value of the fMRI compared to TMS regions was not surprising

The lower predictive value of the fMRI regions compared to the TMS regions is not unexpected. This is because disruptive TMS is used to probe brain function by delivering magnetic pulses that interfere with neural activity in the stimulated region, thereby causing a transient “virtual” lesion (Pascual-Leone et al., 1999, 2000; Walsh and Rushworth, 1999; Walsh and Cowey, 2000). Conversely, fMRI detects regions that are activated by a given task but does not indicate which regions (if any) are necessary for task performance (Price et al., 1999; Rorden and Karnath, 2004; Poldrack, 2008). In short, the causal effect of brain damage on behaviour is more likely to be predicted by TMS regions (that are known to be required for efficient function) than fMRI regions that

might not necessarily be related to the function being investigated. Furthermore, it is possible that the spatial localisation of some of the fMRI effects might have been biased towards the vascular origin of the fMRI signal and away from its neural source (Menon, 2012). This would suggest that TMS-based functional localisation is effective not only because it takes into account inter-subject variability in functional anatomy but also because it is not affected by the vascular architecture of the brain, thereby providing lesion studies with a more accurate starting point.

There are two other factors that could help explain, at least in part, why the predictive power of the fMRI regions was lower than that of the TMS regions. First, the fMRI regions were selected from studies that used phonological tasks that were not exactly the same as those utilised in (i) the TMS studies that identified the TMS regions or (ii) the CAT assessment administered to the stroke patients reported here. The fMRI tasks may, therefore, have engaged different types of processing (e.g., phonological recoding versus covert articulation). Although I examined the degree to which the TMS and CAT phonological tasks were equated, I did not have the data available to assess the degree to which the fMRI and CAT tasks tapped equivalent levels of processing. Second, the fMRI regions were centred around the MNI coordinates from several different studies that each reported distinct parts of SMG [-57, -21, 39; -42, -40, 46; -54, -36, 40] or pOp [-58, 5, 13; -50, 6, 24; -57, 9, 24; -41, 3, 20]. Given inconsistency in the location of effects across multiple studies, and the use of relatively low statistical thresholds, it is plausible that some of the fMRI results were false positives (e.g., Eklund et al., 2016; Chen et al., 2018). Notably, the classification accuracy was not improved when I centred the fMRI regions on the across-study mean coordinates (as opposed to the study-specific coordinates).

Despite the relative advantage of TMS over fMRI in guiding the search for critical lesion sites, it is nonetheless true that fMRI studies have to a greater or lesser degree converged on a set of regions that include those that are necessary for normal phonological performance. Moreover, it remains to be tested whether, in contexts other

than the one investigated here (e.g., when looking at other language or cognitive functions), fMRI-derived regions may prove to be as predictive as (or more than) TMS-derived regions. And, even if this was not the case, fMRI studies would still be essential for pointing TMS studies to parts of the brain that have been associated with normal performance. In other words, fMRI would at the very least provide a rough estimate of the location of critical regions, which can then be further refined by using TMS-based functional localisation.

3.5.4 Explaining inter-patient variability in the effect of damage to the same regions

Although the TMS sites were significantly more predictive than the fMRI sites, the incidence of phonological impairments after damage to the TMS regions (i.e. the positive predictive value) was still below 75%. This across-subject inconsistency in the mapping between lesion and impairment might reflect inter-subject variability in (i) the cognitive strategy used for task performance prior to the stroke (e.g., Seghier et al., 2008b; Kherif et al., 2009; Woollams et al., 2017); (ii) the ability to recover phonological processing skills after damage to the TMS regions (e.g., Specht et al., 2009; Abel et al., 2015; Griffis et al., 2017a; Skipper-Kallal et al., 2017a, b); and/or (iii) the brain regions that support phonological processing (e.g., Price and Friston, 2002; Noppeney et al., 2006; Seghier et al., 2012; Blank et al., 2017). These factors would induce inconsistency in the effect of both real (stroke) and virtual (TMS) lesions.

Additional variability between real and virtual lesions might arise because the sensitivity of the CAT phonological measure (based on the accuracy of non-word reading and digit span) was lower than that used in the TMS studies (i.e. response times to phonological decisions tasks); and/or if the spheres of interest that I used for the lesion study do not match exactly the full extent of grey and white matter that is affected by disruptive TMS.

Another possibility is that persistent (as opposed to transient) phonological impairments might be the consequence of more extensive damage that incorporates the grey and white matter around the original TMS regions. Consistent with this, careful

examination of the lesion sites highlighted that all the patients with the impairment of interest and above-threshold damage to the either of the TMS regions (n = 12 for SMG and n = 13 for pOp) had lesions that included the grey and white matter surrounding the spherical regions of interest. In other words, there is no evidence to conclude that focal damage to either the TMS SMG or pOp regions (i.e. sparing other neighbouring areas) is sufficient to cause persistent phonological impairments. The next chapter (i.e. Experiment II), therefore, pursues how more predictive regions can be selected.

3.5.5 Limitations

There are a number of limitations to the present study that are worth discussing further. First, the tasks used to measure phonological processing abilities in the lesion study (i.e. non-word reading and digit span) were not the same as those utilised in the TMS studies (i.e. phonological decisions; Gough et al., 2005; Sliwinska et al., 2015). This could have artificially reduced the consistency of the effect of "real" and "virtual" lesions on phonological performance, thereby compromising the predictive value of the TMS regions of interest (i.e. it might have inflated the prediction error). Reassuringly, however, (i) non-word reading and digit span have been widely used to index phonological processing (i.e. good face and construct validity) (e.g., Vallar et al., 1997; Silveri and Cappa, 2003; Rapcsak et al., 2009); (ii) neurologically-normal individuals performing these tasks reliably activate SMG and pOp (e.g., Crottaz-Herbette et al., 2004; Wimmer et al., 2010); and (iii) impairments on non-word reading and digit span are associated with damage to SMG and pOp (e.g., Philipose et al., 2007; Baldo et al., 2012; Hickok et al., 2014). Furthermore, the scores from the phonological measure used in the current study (i.e. a combination of non-word reading and digit span) were shown to correlate well with those obtained from the TMS phonological measure (i.e. good convergent validity). And, the scores from the combined non-word reading and digit span measure were significantly more correlated with scores from the TMS phonological measure than scores from the TMS semantic measure (i.e. good discriminant validity). Finally, the results I have presented demonstrate that regions associated with the TMS phonological

measure predicted performance on the CAT phonological measures. This suggests that the function of the regions being investigated is involved in the computation of phonological processes that are shared by a range of language tasks rather than being specific to the TMS phonological task.

Second, in contrast to TMS studies of neurologically-normal subjects where a region's contribution to task efficiency is typically inferred from slowed reaction times during disruptive stimulation (e.g., Gough et al., 2005; Sliwinska et al., 2015), here the effect of damage to the TMS regions on phonological processing was recorded in terms of a persistent impairment of accuracy. This will be less sensitive to phonological processing impairments than response time measurements. However, the effect of damage on accuracy shows how a transient behavioural effect (lasting from milliseconds to minutes) induced by disruptive TMS over SMG or pOp translated into a persistent loss of accuracy after stroke lesions to the same regions.

Third, whereas Gough et al. (2005) and Sliwinska et al. (2015) utilised TMS-based functional localisation to select subject-specific testing sites, the TMS regions of interest used in the current study were centred on the mean MNI stimulation coordinates (i.e. averaged over subjects in the TMS studies). Arguably, this could have resulted in greater localisation error in the lesion study than the TMS studies. However, the observation that phonological processing impairments (as defined here) occurred frequently following stroke damage to the TMS regions suggests that the effect of interest generalized well across studies in spite of these differences.

In summary, the consistency of "real" lesion effects might be even higher if all sources of between-study variation described above are controlled in future experiments.

3.6 Conclusions

This study investigated the accuracy with which damage to regions of interest derived from previous TMS and fMRI investigations of phonological processing in neurologically-normal individuals can predict the incidence of persistent phonological

impairments in stroke patients. It was found that the incidence of phonological impairments was predicted significantly better by the presence or absence of damage to the TMS than fMRI regions. Moreover, the discriminatory power of the TMS sites was not improved further by adding the fMRI sites. Critically, the fact that damage to the TMS regions was reliably associated with the presence of phonological impairments suggests that most patients were not able to fully substitute the function played by TMS SMG and TMS pOp even after years of recovery time, implying that these regions are likely to play a critical role in normal phonological processing. Together, these findings indicate that the use of regions of interest centred on stimulation sites identified with TMS-based functional localisation might, in the future, help us improve our ability to predict language outcome and recovery after stroke.

CHAPTER 4 (Experiment II):

Improving the Prediction Accuracy of the TMS Regions

4.1 Summary

Experiment I showed that the incidence of persistent phonological impairments was predicted significantly better by the presence or absence of damage to regions derived from previous TMS than fMRI studies of phonological processing in neurologically-normal subjects. Critically, however, less than 75% of the patients with above-threshold damage to the TMS regions had phonological impairments; and phonological impairments were observed in 21% of the patients with below-threshold damage to the TMS regions. Experiment II, therefore, sought to determine whether the classification/prediction accuracy of phonological processing impairments could be further improved by adapting the TMS regions to include the surrounding grey and white matter where the presence/absence of damage was consistently associated with the presence/absence of phonological processing impairments.

This novel TMS-guided lesion-deficit mapping approach improved the classification/prediction accuracy, even in a completely independent sample of patients who were more than 5 years post-stroke. Specifically, the odds ratio for the “TMS-guided” regions was more than 3 times higher than that for the original TMS regions. In other words, the ability to discriminate patients with phonological impairments from those without phonological impairments was substantially better when using the TMS-guided regions than the original TMS regions. Likewise, the regions identified using TMS-guided lesion searches yielded a classification accuracy that was approximately 2 times higher than that for regions identified by adapting the borders of the original fMRI regions. These results provide details about the functional anatomy of phonological processing that are not readily available from TMS or fMRI studies in combination or in isolation. For example, the lesion analysis has shown that the critical regions underlying persistent phonological impairments include a combination of grey and white matter in either the vicinity of the left pars opercularis (pOp) or the left supramarginal gyrus (SMG). The findings from Experiment II therefore offer information that is useful for future clinical predictions as well as the interpretation of TMS and fMRI studies of normal phonological processing.

4.2 Introduction

In Experiment I, 26% of the patients with damage to the TMS regions did not have phonological processing impairments. One reason for this may be that the critical lesion sites that cause persistent phonological impairments extend beyond the boundaries of the spherical regions of interest centred on the TMS sites and incorporate the surrounding grey and white matter. The goal of my second study was, therefore, to investigate whether the classification/prediction accuracy can be improved by adapting the borders of the regions of interest to include the grey and white matter around the TMS sites where there is a consistent association between the presence/absence of damage and the presence/absence of phonological impairments. It was hypothesised that adapting the regions of interest to account for the real extent of the lesions in those with phonological impairments should result in an improvement in the classification accuracy. Alternatively, redefining the borders of the regions of interest was not expected to improve classification accuracy if the areas around the TMS sites were co-incidentally damaged because they share the same vasculature as the original TMS regions (Kimberg et al., 2007; Inoue et al., 2014; Mah et al., 2014; Sperber and Karnath, 2017) but are not essential for phonological processing.

Previous lesion-deficit mapping studies in which stroke patients were required to perform phonological decision tasks similar to those utilised to identify the TMS regions (in Gough et al., 2005 and Sliwinska et al., 2015) have already shown that other brain regions lying in close proximity to SMG and pOp, including the white matter underneath, are likely to play an important role in normal phonological processing (Baldo and Dronkers, 2006; Geva et al., 2011; Pillay et al., 2014). For instance, Geva et al. (2011) tested a relatively small sample of 18 stroke patients on two phonological tasks where on each trial they had to decide whether two written words rhymed (i.e. visual rhyme judgements) or whether two written words sounded the same (i.e. visual homophone judgements). Impaired performance on both these phonological decision tasks was associated with damage to a region including the left pars opercularis, the anterior part

of the supramarginal gyrus and underlying white matter, even after factoring out confounds from lesion size or overt speech production. Similarly, another lesion-deficit mapping study (Pillay et al., 2014) assessed 40 chronic left-hemisphere stroke patients on a visual two-choice rhyme matching task in addition to a set of control tasks, and found a significant cluster extending from the left supramarginal gyrus into the posterior part of the left superior temporal gyrus and planum temporale (including the surrounding white matter) where damage was associated with impaired performance, even after removing variance explained by lesion size and overt speech articulation. Furthermore, a wide range of studies involving various experimental procedures and clinical populations have also suggested a role for neighbouring regions in phonological processing, such as the left ventral premotor cortex, posterior part of the left superior temporal gyrus and the white matter tracts running between frontal and parietal sites (e.g., Leff et al., 2009; Chang et al., 2010; Rolheiser et al., 2011; Baldo et al., 2012; Schwartz et al., 2012; Kümmerer et al., 2013; Butler et al., 2014; Cogan et al., 2014; Mandelli et al., 2014; Tate et al., 2014; Henry et al., 2016; Leonard et al., 2016).

A similar picture emerges when considering the findings from functional neuroimaging studies that have shown increased activation in the left supramarginal gyrus and/or pars opercularis of the left inferior frontal gyrus as well as adjacent regions for tasks requiring phonological processing relative to tasks requiring semantic processing (Price et al., 1997; Poldrack et al., 1999; Booth et al., 2002; Devlin et al., 2003; McDermott et al., 2003; Seghier et al., 2004; Gitelman et al., 2005; Simard et al., 2013). For instance, McDermott at al. (2003) had a group of 20 neurologically-normal subjects passively view lists of phonologically-related (e.g., weep, beep, heap) and semantically-related (e.g., bed, rest, awake) words presented in randomly ordered blocks. By contrasting the hemodynamic changes during the phonological and semantic conditions, they concluded that preferential activation for phonological processing was seen in a bilaterally distributed cluster including but not limited to the left premotor cortex, opercular part of the left inferior frontal gyrus and left supramarginal gyrus. Likewise,

Devlin et al. (2003) reported activation in the left supramarginal gyrus, opercular part of the left inferior frontal gyrus and left premotor cortex (among other regions) during an fMRI study of 12 neurologically-normal individuals that compared (i) phonological decisions on whether a word had two syllables and (ii) semantic decisions on whether a word corresponded to a man-made object. Another fMRI study by Seghier et al. (2004) also found greater activation in the left superior temporal gyrus for phonological (do two words rhyme?) relative to semantic (do two words belong to the same semantic category?) decisions, in addition to the left supramarginal gyrus and other areas.

The implication that phonological processing involves regions that neighbour SMG and pOp (such as the left ventral premotor cortex and left posterior superior temporal gyrus) has also been supported by a large number of fMRI studies that have used a wide range of tasks and methods to probe the various components of the phonological network (e.g., Gold et al., 2005; Graves et al., 2008; Papoutsi et al., 2009; McGettigan et al., 2011; Chevillet et al., 2013; Hope et al., 2014; Mei et al., 2014; Evans and Davis, 2015; Goranskaya et al., 2016). Moreover, TMS delivered directly over the left premotor cortex (Meister et al., 2007; Sato et al., 2009; Krieger-Redwood et al., 2013) or the posterior part of the left superior temporal gyrus (Acheson et al., 2011), for example, has been shown to disrupt the efficiency with which phonological tasks can be performed (Murakami et al., 2015). From a neural network perspective, where multiple brain regions contribute to any one task via dynamic interactions, it follows that the constituents of the phonological system should share dense structural and functional connections. Indeed, this is in line with what is known about the so called dorsal language stream that underpins the mapping between sensory and motor speech codes for overt and covert speech processing (e.g., Hickok and Poeppel, 2007; Saur et al., 2008; Xiang et al., 2010; Pei et al., 2011; Herman et al., 2013; Liebenthal et al., 2013; Margulies and Petrides, 2013; Dick et al., 2014; Fridriksson et al., 2016; Rolston and Chang, 2018).

4.2.1 Research question

In brief, the current experiment attempted to answer the following research question:

- Can the prediction accuracy be improved by adapting the regions of interest assessed in Experiment I to include the grey and with matter areas around the TMS sites where the presence/absence of damage is consistently associated with the presence/absence of phonological processing impairments?

The new regions of interest were generated using data from patients reported in Experiment I (Sample 1). I then compared the classification/prediction accuracy of (a) the new regions and (b) the original TMS regions, in an independent sample of patients who differed from those used in Experiment I (1-5 years post-stroke) because they were tested more than 5 years after their stroke. In addition to comparing the classification accuracy of the new and original TMS regions, I was also able to investigate whether a similar classification performance could be obtained by adapting the borders of regions identified in previous fMRI rather than TMS studies of neurologically-normal individuals.

4.3 Materials and Methods

4.3.1 Participants

There were two groups of participants. Sample 1 included data from the same 154 patients that were selected in Experiment I (see Table 3.1 in Chapter 3). In Experiment II, these participants were used to investigate how the TMS regions of interest from Experiment I could be adapted (into TMS-guided regions) to improve the classification accuracy (i.e. how well the presence or absence of damage predicts the incidence of phonological impairments). The predictive power of the new regions was tested in a second (independent) sample of patients. This independent sample (Sample 3) included 108 participants who had been excluded from Sample 1 (in Experiment I) because they were tested more than 5 years post-stroke onset. The overlap between Sample 3 in the current chapter (Experiment II) and Sample 2 in Experiment I was 16 patients only.

Just like those in Sample 1, the patients in Sample 3 were all adult stroke survivors who: (1) had a left-hemisphere lesion (as attested by a clinical neurologist) that was greater than 1 cm^3 (as measured by an automated lesion identification tool; Seghier et al., 2008a); (2) had no history of neurological or psychiatric illness that were not related to their stroke; (3) were right-handed (pre-morbidly); (4) were native speakers of English (either monolingual or bilingual); and (5) had normal or corrected-to-normal vision and hearing (as per self-reports). See Table 4.1 for demographic and clinical details and Figure 4.1 for lesion overlap map. The prevalence of phonological impairments in Sample 3 (40/108 = 37%) was approximately the same as in Sample 1 (61/154 = 40%).

Table 4.1: Summary of demographic and clinical data for Sample 3.

Factor		**Sample 3 (*N* = 108)**
Age at stroke onset (years)	*M*	50.6
	SD	12.2
	Range	17.2-72.3
Age at scan acquisition (years)	*M*	60.8
	SD	11.0
	Range	33.2-83.6
Time post-stroke (years)	*M*	10.1
	SD	6.0
	Range	5.1-36.0
Lesion size (cm^3)	*M*	129.3
	SD	106.5
	Range	1.2-405.0
Gender	Males	66
	Females	42

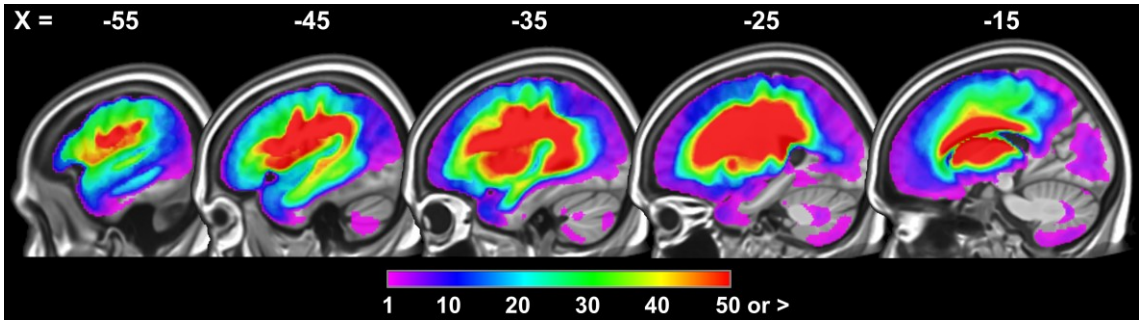

Figure 4.1: Lesion overlap map for Sample 3. Lesion overlap map of the 108 stroke patients included in Sample 3, depicting the frequency of overlapping lesions at each given voxel.

Despite the similarity in the inclusion criteria used to allocate patients to Samples 1 and 3, I expected that prediction accuracy would be less in Sample 3 because these patients had longer to recover (i.e. > 5 years post-stroke) than those in Sample 1 (i.e. 1-5 years post-stroke). However, this was not of concern because the observation of interest is whether the presence or absence of phonological impairments in Sample 3 is better predicted by the TMS-guided regions than the original TMS regions or fMRI-guided regions (i.e. predictions are compared on the same group of patients).

Comparison of demographic and clinical variables showed that, on average, the patients in Sample 3 had (i) significantly larger lesions ($t(188) = 4.06$, $p < 0.001$) and were significantly (ii) younger at the time of stroke onset ($t(260) = 3.57$, $p < 0.001$) and (iii) later post-stroke ($t(113) = 12.68$, $p < 0.001$) than the patients in Sample 1. For formal assessment of differences in lesion distributions and CAT scores, see Figure 4.2 and Table 4.2, respectively.

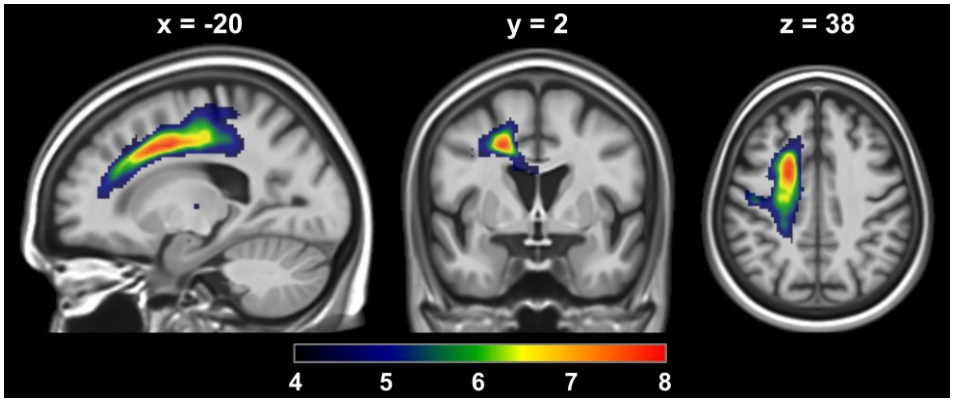

Figure 4.2: Comparison of the distribution of lesions between Sample 1 and Sample 3. The figure depicts areas of the brain where the degree of damage was significantly greater in patients from Sample 3 than Sample 1. No brain regions were identified where the degree of damage was significantly greater in patients from Sample 1 than Sample 3. The statistical map comprises voxels that survived a voxel-level family-wise error (FWE) corrected threshold (i.e. $p < 0.05$, FWE-corrected) as well as an extent threshold for each cluster of 20 voxels. The colour scale shows the range of t values associated with statistically significant effects.

4.3.2 Defining phonological processing impairments

As in Experiment I, the patients were classified as having "phonological impairments" when their performance on both the non-word reading and digit span tasks from the CAT was within the impaired range (for rationale, see Chapter 3). In addition,

the severity of phonological impairments was measured by averaging scores across the non-word reading and digit span tasks.

Table 4.2: Comparison of CAT scores between Sample 1 and Sample 3.

CAT Task No.	Sample 1 *M*(±*SD*)	Sample 3 *M*(±*SD*)	*t*	*df*	*p*
1	55.0(±9.0)	55.0(±8.7)	0.07	246	0.946
2	56.7(±5.8)	56.9(±6.4)	0.33	254	0.739
3	57.9(±10.7)	59.1(±10.0)	0.97	259	0.333
4	53.7(±7.4)	53.8(±6.3)	0.18	254	0.857
5	58.9(±7.9)	60.4(±8.3)	1.48	253	0.139
6	57.4(±7.4)	57.6(±6.7)	0.18	253	0.860
7	56.4(±6.9)	57.5(±6.5)	1.26	260	0.210
8	55.8(±7.7)	56.8(±7.1)	0.97	260	0.332
9	57.2(±8.7)	58.5(±8.1)	1.23	260	0.220
10	57.7(±9.1)	59.1(±8.2)	1.29	258	0.199
11	53.6(±8.1)	54.7(±7.7)	1.08	258	0.280
12	54.0(±8.8)	56.7(±8.7)	2.47	260	0.014
13	53.3(±10.0)	56.4(±8.7)	2.62	249	0.009
14	53.5(±8.8)	56.7(±9.5)	2.75	260	0.006
15	52.9(±9.1)	53.9(±9.0)	0.95	260	0.344
16	54.2(±9.4)	55.0(±8.3)	0.80	247	0.424
17	57.5(±11.0)	60.9(±10.1)	2.51	260	0.013
18	55.9(±10.2)	57.2(±9.8)	1.05	259	0.293
19	55.3(±9.5)	56.9(±8.7)	1.36	255	0.175
20	55.2(±9.6)	58.3(±9.6)	2.55	260	0.011
21	54.0(±11.4)	56.5(±11.2)	1.72	259	0.087
22	55.0(±9.8)	57.6(±7.9)	2.38	255	0.018
23	53.6(±10.7)	55.4(±11.4)	1.26	260	0.207
24	58.1(±6.0)	58.5(±5.2)	0.62	259	0.534
25	57.9(±9.0)	59.3(±8.3)	1.24	257	0.218
26	56.6(±9.1)	57.1(±8.7)	0.42	255	0.672
27	60.5(±9.8)	61.9(±9.2)	1.09	249	0.279

None of the *p* values (right-most column) survived a Bonferroni correction for multiple comparisons (i.e. $p < 0.002$). See Chapter 2 for which tasks correspond to each number. *t* = t statistic; *df* = degrees of freedom.

4.3.3 Generating new TMS-guided and fMRI-guided regions

Using a subset of data from Sample 1 only, the borders of the TMS regions were adapted to take into account the extent of damage in those with phonological impairments. To do this, I created lesion overlap maps for patients with phonological impairments who either had damage to (1) the TMS SMG but not pOp regions; or (2) the

TMS pOp but not SMG regions. The brain areas where the presence or absence of damage best explained the presence or absence of phonological impairments became the TMS-guided regions.

The fMRI-guided regions were identified in the same way as the TMS-guided regions, with the only exception that patients (from Sample 1) with damage to spherical regions of interest centred on the mean fMRI coordinates ([-52, 6, 20] for pOp and [-51, -32, 42] for SMG) were used to create the lesion overlap maps.

4.3.4 Determining the threshold for critical damage

To determine the degree of damage that maximised classification accuracy, I systematically investigated how well patients with and without phonological impairments were segregated when they either had 70%, 80%, 90% or 100% damage to the TMS-guided or fMRI-guided regions. This was achieved by selecting two subsets of patients from Sample 1 (with and without phonological impairments) that were matched for lesion size (i.e. Subset 1A and Subset 1B). Subset 1A included patients who had selective phonological processing impairments (i.e. phonological but not semantic impairments; see Chapter 3). Subset 1B included those who did not meet the criteria used to define phonological impairments. Lesion size was matched across Subsets 1A and 1B by finding the minimum and maximum lesion volumes that were common to both groups, with no significant differences in mean lesion size across groups (see Table 3.4 in Chapter 3).

For each percentage of damage (70%, 80%, 90% and 100%), the classification accuracy for discriminating between Subset 1A (with phonological impairments) and Subset 1B (without phonological impairments) was expressed in terms of positive and negative predictive values, sensitivity, specificity and the odds ratio (Altman and Bland, 1994a, b; Bland and Altman, 2000; Glas et al., 2003). The threshold was set at the degree of damage that had the highest odds ratio, which quantitatively describes the strength of the association between the presence/absence of a deficit of interest (i.e. phonological impairments) and the presence/absence of damage to a particular set of

regions. The generalizability of these thresholds was then tested in an independent sample of patients (the validation group described above).

4.4 Results

4.4.1 The TMS-guided regions

The new TMS-guided regions incorporated the grey and white matter, around the original TMS sites (i.e. 0.5 cm^3 spheres), that was consistently damaged in the lesion overlap maps (see Figures 4.3A and 4.4A).

For pOp, the new region was 20.6 cm^3 in size and extended deep into the underlying white matter, including portions of the anterior and long segments of the arcuate fasciculus, uncinate fasciculus, inferior fronto-occipital fasciculus, external and internal capsule as well as portions of the left ventral premotor cortex and caudate nucleus (see Figure 4.3B). This TMS-guided pOp region was 100% damaged in 12/13 of the patients with above-threshold damage to the original TMS pOp region and phonological impairments. Of note, 80% of the original TMS pOp regions was contained within the new TMS-guided pOp region (see Table 4.3).

For SMG, the new region was 24.9 cm^3 in size and extended deep into the underlying white matter, including portions of the anterior, posterior and long segments of the arcuate fasciculus, the inferior longitudinal fasciculus and inferior fronto-occipital fasciculus as well as portions of the posterior part of the superior temporal gyrus (see Figure 4.4B). This TMS-guided SMG region was at least 85% damaged in 11/12 of the patients with above-threshold damage to the original TMS SMG region and phonological impairments. Of note, 82% of the original TMS SMG regions was contained within the new TMS-guided SMG region (see Table 4.3).

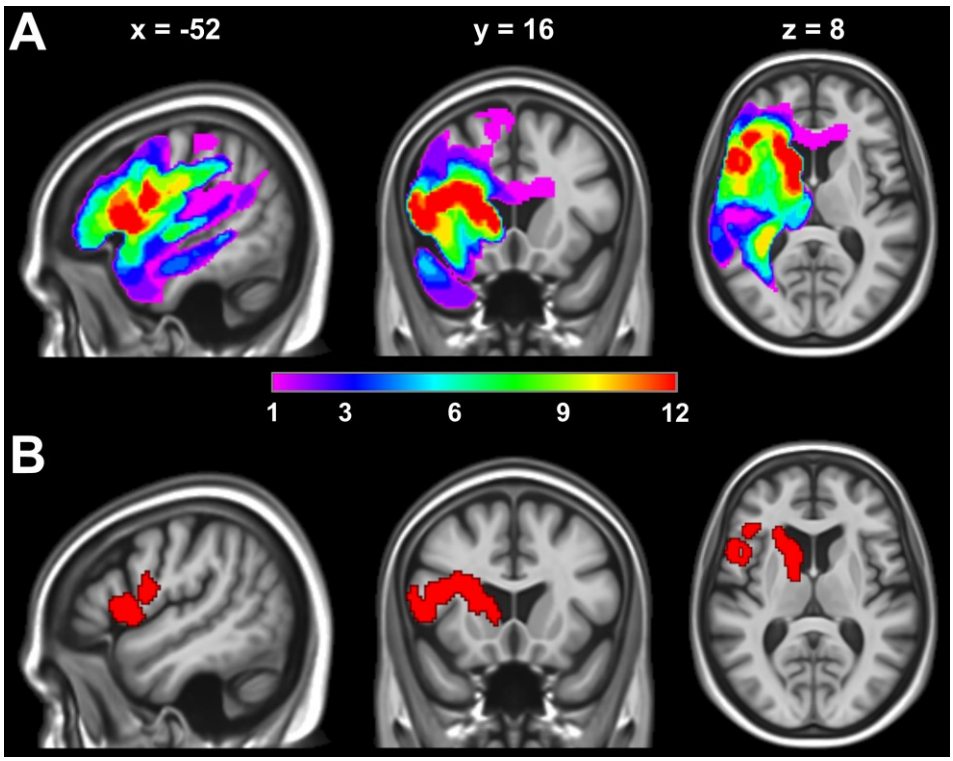

Figure 4.3: TMS-guided pOp region. (A) Lesion overlap map of patients with phonological impairments following above-threshold damage to the TMS pOp region (*n* = 12). (B) The TMS-guided pOp region (in red) only included voxels that were damaged in 12 out of 12 patients from the lesion overlap map.

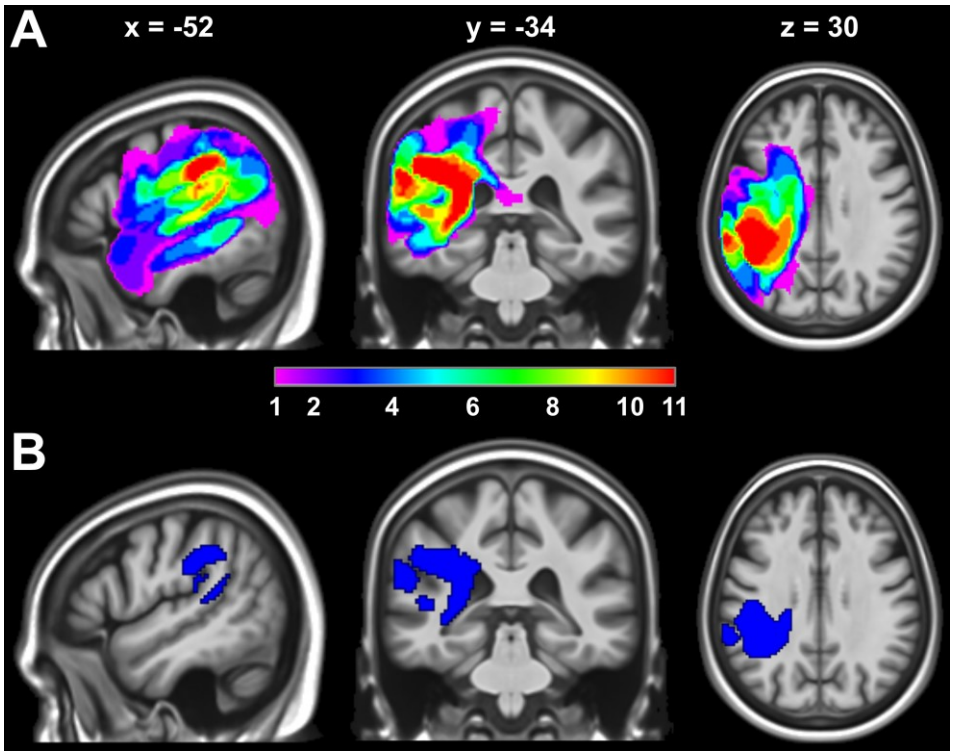

Figure 4.4: TMS-guided SMG region. (A) Lesion overlap map of patients with phonological impairments following above-threshold damage to the TMS SMG region (*n* = 11). (B) The TMS-guided SMG region (in blue) only included voxels that were damaged in at least 10 out of 11 patients from the lesion overlap map.

4.4.2 The fMRI-guided regions

The new fMRI-guided regions included the grey and white matter, around the original fMRI sites (i.e. 0.5 cm^3 spheres centred on the across-study mean coordinates), that was consistently damaged in the lesion overlap maps (see Figure 4.5A).

For pOp, the new region was 18.1 cm^3 and encompassed portions of the ventral sensorimotor cortex, ventral premotor cortex and short gyri of the insula as well as portions of the anterior and long segments of the arcuate fasciculus (see Figure 4.5B). This fMRI-guided pOp region was 100% damaged in 17/17 of the patients with above-threshold damage to the original fMRI pOp region and phonological impairments.

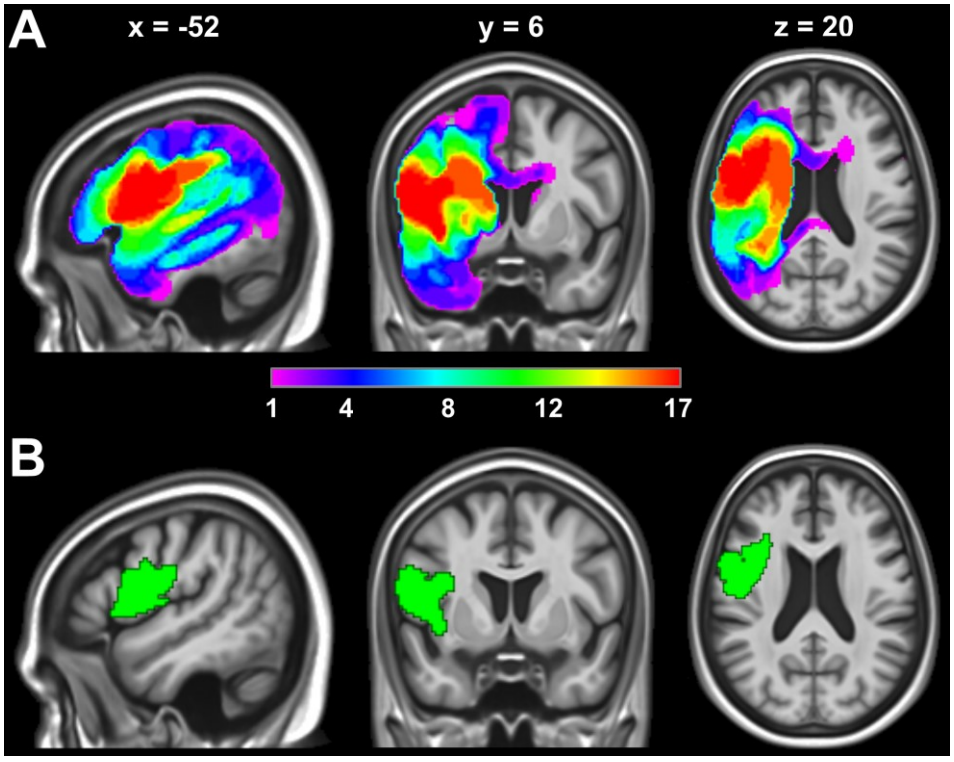

Figure 4.5: fMRI-guided pOp region. (A) Lesion overlap map of patients with phonological impairments following above-threshold damage to the mean fMRI pOp region (*n* = 17). (B) The fMRI-guided pOp region (in green) only included voxels that were damaged in 17 out of 17 patients from the lesion overlap map.

For SMG, the new region was based on one patient only, because no other patients with phonological impairments had selective damage to the original fMRI SMG region while sparing the fMRI pOp region. Consequently, the new fMRI-guided SMG region was very large (161.8 cm^3) and encompassed most of the left perisylvian cortex (see Figure 4.6), including 94% of the fMRI-guided pOp region (see Table 4.3).

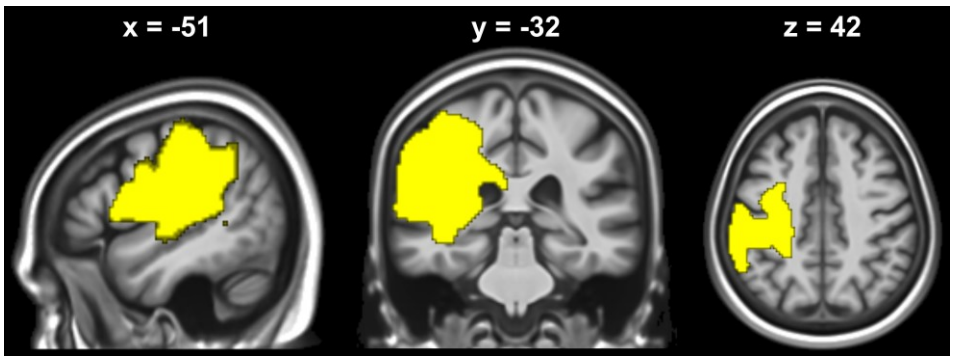

Figure 4.6: fMRI-guided SMG region. The lesion of one patient only (i.e. PS0288) contributed to the fMRI-guided SMG region, because no other patients had phonological impairments in the presence of above-threshold damage to the mean fMRI SMG region and below-threshold damage to the mean fMRI pOp region.

Table 4.3: Overlap between TMS, TMS-guided and fMRI-guided regions.

		TMS		TMS-guided		fMRI-guided	
		SMG	**pOp**	**SMG**	**pOp**	**SMG**	**pOp**
TMS	**SMG**	100%	0%	3%	0%	1%	0%
	pOp	0%	100%	0%	4%	0%	3%
TMS-guided	**SMG**	82%	0%	100%	0%	15%	1%
	pOp	0%	80%	0%	100%	11%	39%
fMRI-guided	**SMG**	100%	46%	100%	90%	100%	94%
	pOp	0%	50%	1%	35%	11%	100%

The table shows the percentage of each of the regions listed at the top contained within each of the regions on the left-hand side.

4.4.3 Critical damage thresholds

The degree of damage that best segregated patients with selective phonological impairments (Subset 1A) from those without phonological impairments (Subset 1B) was 80% for TMS-guided regions and 90% for fMRI-guided regions (see Table 4.4). Henceforth, “above-threshold damage” refers to instances where the degree of damage to the region of interest was equal to or greater than these thresholds.

Table 4.4: Critical damage thresholds for TMS-guided and fMRI-guided regions.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
TMS-guided	100%	70%	64%	30%	91%	4.2
	90%	69%	67%	39%	88%	4.5
	80%	70%	74%	61%	81%	***6.7***
	70%	54%	70%	65%	59%	2.7
fMRI-guided	100%	100%	60%	9%	100%	---
	90%	67%	61%	17%	94%	***3.2***
	80%	57%	60%	17%	91%	2.0
	70%	55%	61%	26%	84%	1.9

Threshold = percentage of the region of interest (ROI) damaged; PPV/NPV = positive and negative predictive values.

4.4.5 Classification accuracy for TMS-guided regions in Sample 1

Over all 154 patients in Sample 1, the incidence of phonological impairments was 85% (46/54) in those with above-threshold damage to one or both of the TMS-guided regions and 15% (15/100) in those with below-threshold damage to both TMS-guided regions. When the classification accuracy of the TMS-guided regions was re-expressed in terms of the odds ratio (i.e. a single metric that reflects the overall performance of a binary predictor/classifier), it was observed that the odds for the presence of phonological impairments in those with above-threshold damage to the TMS-guided regions was 32.6 times higher than the odds for the presence of phonological impairments in those with below-threshold damage to the TMS-guided regions. This basically means that the discriminatory ability (or power) of the TMS-guided regions was very high.

When the effect of each lesion site was considered separately, it was found that the incidence of phonological impairments after damage to TMS-guided “SMG and pOp”, “SMG not pOp”, and “pOp not SMG” was: 89% (17/19), 90% (9/10) and 80% (20/25), respectively. In those with below-threshold damage to both regions, the incidence of phonological impairments was 5% (2/39) in the context of 0-20% damage to SMG and/or pOp, rising to 21% (13/61) in those with 21-79% damage to either SMG or pOp or both.

Critically, the effect of damage to the TMS-guided regions cannot be explained by lesion size because, even after matching lesion size across groups (see Tables 4.5

and 4.6), the incidence and severity of phonological processing impairments were significantly worse in patients with above-threshold damage to either of the TMS-guided regions than in control patients (i.e. those with below-threshold damage to both TMS-guided regions).

Table 4.5: Summary of demographic and clinical data for lesion-size-matched groups.

Factor		**SMG** ***n* = 8**	**pOp** ***n* = 13**	**Control** ***n* = 30**
Age at stroke onset (years)	*M*	57.4	48.3	57.6
	SD	10.2	14.7	12.2
	Range	44.2-76.9	27.8-72.9	24.8-75.4
Age at scan acquisition (years)	*M*	59.9	50.6	60.3
	SD	10.2	14.5	12.0
	Range	46.5-78.3	30.7-74.1	29.4-78.2
Time post-stroke (years)	*M*	2.5	2.3	2.7
	SD	1.3	0.9	1.0
	Range	1.1-4.5	1.1-3.8	1.0-4.6
Lesion size (cm^3)	*M*	85.1	90.8	77.6
	SD	29.2	20.9	19.8
	Range	44.3-117.8	57.6-128.7	51.5-127.0
Gender	Males	7	9	26
	Females	1	4	4
Phonological performance	Imp (Not)	8 (0)	10 (3)	10 (20)
	M	45.1	46.9	53.1
	SD	4.2	7.0	7.3
	Range	37.5-50.0	37.5-60.0	37.5-67.0
% damage to TMS-guided SMG	*M*	93.7	30.8	46.8
	SD	6.1	27.3	26.2
	Range	86.0-100.0	0.0-72.0	0.0-79.0
% damage to TMS-guided pOp	*M*	19.8	93.3	33.2
	SD	23.9	8.2	28.3
	Range	1.0-60.0	80.0-100.0	0.0-78.0

The three patient groups listed at the top of the table were matched for lesion size; $F(2,48) = 1.78$, $p = 0.180$. Imp = number of patients with phonological impairments; Not = number of patients who did not meet the criteria for phonological impairments.

Table 4.6: Comparison of the incidence and severity of phonological processing impairments.

	Incidence		Severity	
Group	SMG vs. Control	pOp vs. Control	SMG vs. Control	pOp vs. Control
n	8 vs. 30	13 vs. 30	8 vs. 30	13 vs. 30
% / M	100% vs. 33%	77% vs. 33%	45.1 vs. 53.1	46.9 vs. 53.1
X^2 / t	[a]	6.93	2.91	2.56
df	[a]	1	36	41
OR / d	[a]	6.67	1.16	0.85
p	0.001	0.008	0.006	0.014

[a] = Fisher's exact test; % = percentage of patients with phonological impairments from the corresponding patient group; *M* = mean phonological score for each patient group; X^2 = chi-square statistic; *t* = t statistic; *df* = degrees of freedom; *OR* = odds ratio; *d* = Cohen's d.

4.4.6 Comparing the classification accuracy for TMS and TMS-guided regions

4.4.6.1 Sample 1

Over all 154 patients in Sample 1, the incidence of phonological impairments rose to 85% (46/54) from 74% (40/54) in those with above-threshold damage to the TMS-guided versus TMS regions while falling to 15% (15/100) from 21% (21/100) in those with below-threshold damage to the TMS-guided versus TMS regions (see Table 4.7 and Figure 4.7); which resulted in a higher classification accuracy for the TMS-guided regions relative to the original TMS regions (odds ratio = 32.6 versus 10.7). However, this is not surprising given that the TMS-guided regions were defined from the lesion sites of 25 patients who had impaired phonological processing after damage to one or the other of the TMS regions. Data from these 25 patients were therefore excluded to compare how well the new regions explained outcome in other patients in Sample 1 (i.e. cross-validation). This was only possible in 10 patients who had below-threshold damage to both TMS sites but above-threshold damage to the TMS-guided sites (one SMG, eight pOp and one both); see Table 4.8. As 8/10 of these patients had phonological impairments, outcome was 80% consistent with the predictions of the TMS-guided regions (where above-threshold damage predicts the presence of phonological

impairments) but only 20% consistent with the predictions of the original TMS regions (where below-threshold damage predicts the absence of phonological impairments).

Table 4.7: Classification accuracy for TMS, TMS-guided and fMRI-guided regions in Sample 1.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
(1) TMS	80%	74%	79%	66%	85%	10.7
(2) TMS-guided	80%	85%	85%	75%	91%	32.6
(3) fMRI-guided	90%	88%	73%	46%	96%	18.9
(2) & (3)	as above	85%	86%	77%	91%	35.7

Threshold = critical damage threshold for each region of interest (ROI); PPV/NPV = positive and negative predictive values.

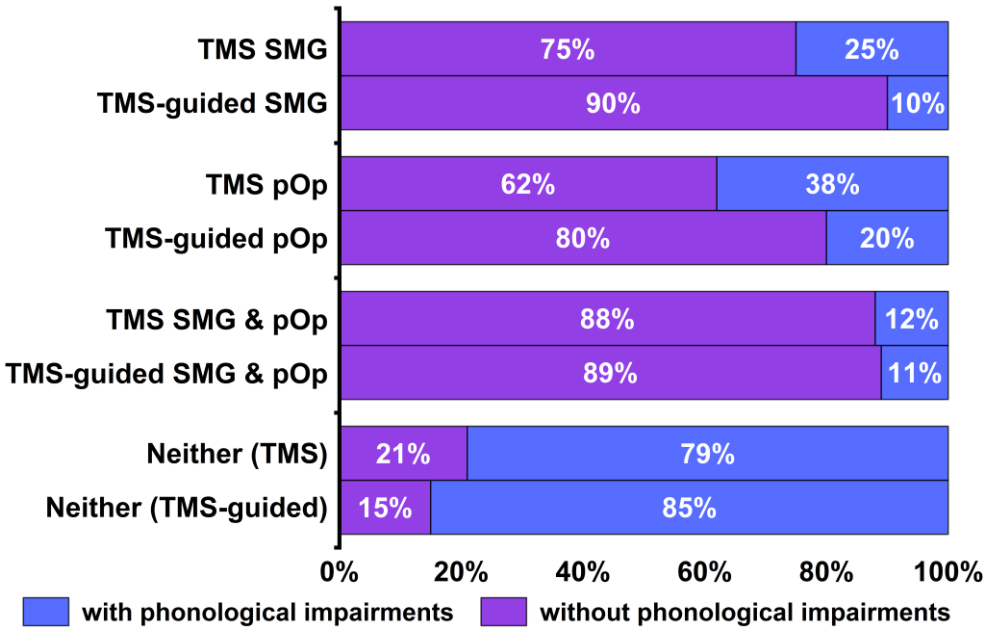

Figure 4.7: Classification accuracy for TMS and TMS-guided regions. Improvements in the classification accuracy can be seen when the lesion categorisation changed from original TMS regions to TMS-guided regions. Patients with above-threshold damage to the TMS-guided regions exhibited higher incidence of phonological impairments (i.e. impaired non-word reading and digit span) than patients with above-threshold damage to the original TMS regions.

Table 4.8: Lesion categorisation.

Above-threshold damage to:	TMS-guided SMG	TMS-guided pOp	TMS-guided SMG & pOp	Neither
TMS SMG	9 (8)	0	3 (3)	4 (1)
TMS pOp	0	15 (12)	0	6 (1)
TMS SMG & pOp	0	2 (2)	15 (13)	0
Neither	1 (1)	8 (6)	1 (1)	90 (13)

The numbers of patients who moved from one lesion group to another, with number of patients who had phonological processing impairments (i.e. impaired performance on both non-word reading and digit span) shown in parentheses.

4.4.6.2 Sample 3

In an independent sample of patients (i.e. patients who were not used to define the regions), the classification accuracy of the TMS-guided regions was more than 3 times higher than that of the original TMS regions (odds ratio = 27.6 versus 8.2). In other words, the ability to discriminate between those with and without phonological impairments was substantially better when using the TMS-guided regions relative to using the original TMS regions.

As can be seen from Table 4.9, the improvement in the classification accuracy was primarily because the TMS-guided regions were better able to explain the absence of phonological impairments (i.e. the negative predictions). Specifically, below-threshold damage resulted in phonological impairments in only 6% of patients (i.e. the error was very low), but error was higher (18%) in those who had below-threshold damage to the original TMS regions. The positive predictions (i.e. of who will have a deficit) were less accurate (perhaps because many patients had recovered). The critical point, however, is that the positive predictive value was roughly the same if predictions were based on the TMS-guided regions (64%) or the original TMS regions (65%). Thus, over positive and negative predictions, the presence of phonological impairments could almost entirely be accounted for by damage to the TMS-guided regions (i.e. 93% compared to 70% for the original TMS regions).

Table 4.9: Classification accuracy for TMS, TMS-guided and fMRI-guided regions in Sample 3.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
(1) TMS	80%	65%	82%	70%	78%	8.2
(2) TMS-guided	80%	64%	94%	93%	69%	27.6
(3) fMRI-guided	90%	65%	76%	55%	82%	5.7
(2) & (3)	as above	64%	94%	93%	69%	27.6

Threshold = critical damage threshold for each region of interest (ROI); PPV/NPV = positive and negative predictive values.

4.4.7 Comparing the classification accuracy for fMRI-guided and TMS-guided regions

4.4.7.1 Sample 1

The classification accuracy for the TMS-guided regions was nearly 2 times higher than that for the fMRI-guided regions (odds ratio = 32.6 versus 18.9). This mainly occurred because the TMS-guided regions were better able to explain the absence of phonological impairments (i.e. the negative predictions). Specifically, below-threshold damage to the TMS-guided regions resulted in phonological impairments in 15% of patients compared to 27% of patients with below-threshold damage to the fMRI-guided regions. Consequently, a greater proportion of the patients with phonological impairments was accounted for by above-threshold damage to the TMS-guided (75%) than fMRI-guided (46%) regions (see Table 4.7).

When the SMG regions were not considered, both the TMS-guided and fMRI-guided pOp regions yielded identical classification accuracies (odds ratio = 18.9 for both). This occurred primarily because the positive and negative predictive values of the TMS-guided pOp region and the fMRI-guided pOp region were very similar; see Table 4.10 for details. However, a significantly larger proportion of the patients with phonological impairments were explained by above-threshold damage to the TMS-guided pOp region (61%) than the fMRI-guided pOp region (46%); $X^2 = 7.36$, $p = 0.007$.

Table 4.10: Classification accuracy for TMS-guided and fMRI-guided pOp regions.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
Sample 1 (1-5 years post-stroke)						
TMS-guided pOp	80%	84%	78%	61%	92%	18.9
fMRI-guided pOp	90%	88%	73%	46%	96%	18.9
Sample 3 (> 5 years post-stroke)						
TMS-guided pOp	80%	59%	81%	73%	71%	6.3
fMRI-guided pOp	90%	65%	76%	55%	82%	5.7

Threshold = critical damage threshold for each region of interest (ROI); PPV/NPV = positive and negative predictive values.

4.4.7.2 Sample 3

It was found that, although Sample 3 was entirely independent of the region identification process, the classification accuracy of the TMS-guided regions (odds ratio = 27.6) was more than 4 times higher than that of the fMRI-guided regions (odds ratio = 5.7). As in Sample 1, the improvement in the classification accuracy was primarily because below-threshold damage to the TMS-guided regions predicted the absence of phonological impairments more accurately than below-threshold damage to the fMRI-guided regions (94% versus 76% accuracy); see Table 4.9. Consequently, a substantially greater proportion of the patients with phonological impairments was explained by above-threshold damage to the TMS-guided than fMRI-guided regions (93% versus 55%).

When the SMG regions were not considered, the classification accuracy of the TMS-guided pOp region was similar to that of the fMRI-guided pOp region (odds ratio = 6.3 versus 5.7); see Table 4.10 for details. Again, however, a significantly larger proportion of the patients with phonological processing impairments could be accounted for by above-threshold damage to the TMS-guided pOp region (73%) than the fMRI-guided pOp region (55%); $X^2 = 5.44$, $p = 0.020$.

4.4.8 Inconsistency in the effect of TMS-guided pOp lesions in Sample 3

Consideration of the effect of each lesion site separately, showed that above-threshold damage to the TMS-guided pOp region only inconsistently resulted in

phonological impairments in Sample 3 (i.e. > 5 years post-stroke). Specifically, 13/26 (50%) of the patients with above-threshold damage to TMS-guided pOp had phonological impairments (compared to, for instance, 8/9 = 89% for those with above-threshold damage to TMS-guided SMG). To investigate what factors could explain the inconsistency in the mapping between lesion and deficit for TMS-guided pOp, I generated the hypothesis that the subsets of patients with versus without phonological impairments in the context of above-threshold damage to the TMS-guided pOp region may differ in some critical way (e.g., age at stroke onset or time post-stroke).

By examining the demographics of patients in Sample 3 with phonological impairments (n = 13) versus without phonological impairments (n = 13) in the context of above-threshold damage to the TMS-guided pOp region (but below-threshold damage to TMS-guided SMG), I found that the patients without phonological impairments were (i) tested later post-stroke (M(±SD) = 7.7(±2.5) versus 11.1(±4.4) years, t(19) = 2.40, p = 0.027) and (ii) younger at the time of stroke onset (M(±SD) = 57.3(±8.9) versus 46.7(±12.5) years, t(24) = 2.49, p = 0.020), with (iii) no significant differences in mean lesion size (M(±SD) = 186.0(±90.4) versus 169.7(±94.7) cm^3, p = 0.658). Together, these results seem to be consistent with the possibility that a subset of the patients with damage to TMS-guided pOp might have been able to recover their phonological processing skills after years of recovery time.

Preliminary evidence in favour of the recovery hypothesis comes from a small group of patients whose phonological processing abilities were assessed with the CAT more than once. Specifically, 13 patients from Sample 1 (i.e. 1-5 years post-stroke) with phonological impairments in the context of above-threshold damage to the TMS-guided pOp region were followed up at least one year after the initial assessment session (M = 30.1 months, SD = 13.2, range = 14 to 49 months). Comparison of task scores confirmed that these patients had, on average, experienced a significant improvement in phonological performance between time point 1 and 2 (M = 43.9 versus 45.7; t(12) = 2.03, p = 0.033 one-tailed). Post-hoc analyses revealed that the behavioural changes

were mainly driven by digit span scores (M = 43.9 versus 47.3, p = 0.021 one-tailed), with no significant contribution from non-word reading scores (M = 43.9 versus 44.2, p = 0.426 one-tailed). Overall, this preliminary set of results appears to suggest that inter-patient variability in the recovery of language function might play a role in explaining inconsistency in the mapping between the presence of TMS-guided pOp damage and the presence of phonological impairments.

4.5 Discussion

Experiment II identified new regions of interest that included the grey and white matter that was damaged in patients with phonological impairments and lesions to either the TMS pOp region or the TMS SMG region. The classification accuracy of these new "TMS-guided" regions was found to be substantially better than that of the original TMS regions, even when tested in an independent sample of patients (i.e. Sample 3). Similarly, the classification accuracy of the TMS-guided SMG and pOp regions was substantially better than that of the fMRI-guided SMG and pOp regions irrespective of time post-stroke.

In Sample 1 (1-5 years post-stroke), the TMS-guided regions outperformed the original TMS regions in every single metric of classification accuracy including positive and negative predictive values, sensitivity and specificity. In Sample 3 (> 5 years post-stroke), the advantage of the new TMS-guided regions arose at the level of negative predictions and sensitivity. This is because below-threshold damage to the TMS-guided regions was very consistently associated with spared phonological processing abilities (i.e. high negative predictive value). Put another way, the incidence of phonological impairments was almost entirely explained by above-threshold damage to the TMS-guided regions (i.e. high sensitivity), which means that most (if not all) of the regions that are critical for phonological processing were successfully identified.

The finding that damage to TMS-guided SMG and/or pOp accounts for the majority of patients with phonological impairments has implications for understanding the

functional anatomy of phonological processing and are also relevant for making clinical predictions for future patients.

4.5.1 The effect of damage to the TMS-guided regions on phonological processing

The TMS-guided region for SMG included parts of the left posterior superior temporal gyrus, arcuate fasciculus, inferior longitudinal fasciculus and inferior fronto-occipital fasciculus. Damage to this region caused phonological processing impairments in more than 85% of patients irrespective of the time post-stroke tested (1-5 years in Sample 1 and > 5 years in Sample 3). The TMS-guided region for pOp included parts of the left ventral premotor cortex, caudate nucleus, arcuate fasciculus, uncinate fasciculus, inferior fronto-occipital fasciculus, external and internal capsule. The effect of damage to this region was less consistent than the effect of damage to the TMS-guided SMG region, particularly in patients who were tested more than 5 years after their stroke (i.e. Sample 3). Plausibly, it might be the case that the potential for recovery is greater after damage to the TMS-guided pOp region than the TMS-guided SMG region. Indeed, preliminary longitudinal evidence indicated that phonological performance improved over time despite TMS-guided pOp damage in a small subset of patients from Sample 1 whose language abilities were assessed more than once after stroke.

Prior studies of clinical populations (when considered individually) had already associated subsets of these areas with phonological processing skills (e.g., Abdullaev et al., 1998; Tettamanti et al., 2005; Teichmann et al., 2009; Rolheiser et al., 2011; Geva et al., 2011; Henry et al., 2016; Halai et al., 2017). Over and above what was previously shown, Experiment II revealed the full extent of damage that could cause persistent phonological processing impairments. The extent of the TMS-guided regions also suggests that persistent phonological impairments might be the consequence of a combination of grey and white matter damage. This provides details about the functional anatomy of phonological processing that are not readily available from TMS or fMRI studies together or in isolation because these techniques primarily assess cortical function. The further limitation for TMS studies is that they can only target a few testing

sites located on the cortical surface and therefore tens of experiments would be required to conclude that all cortical regions that are important for task efficiency have been identified. In contrast, although fMRI can measure cortical activity across the whole of the brain, task-related activity may not be essential for task performance or even contribute to task performance (e.g., Price et al., 1999). Conversely, a limitation of the current lesion study is that the data I have reported do not allow me to answer the question of what proportion of the TMS-guided regions is truly critical for the presence of phonological impairments and what proportion of the TMS-guided regions is merely a by-product of co-occurring patterns of stroke damage caused by the vascular anatomy of the brain.

On the other hand, the findings reported in this chapter, could not have been reached without prior fMRI and TMS studies of neurologically-normal individuals. Whole brain fMRI studies were used to pinpoint task-related cortical sites and thereby guide the TMS studies by specifying which cortical regions are activated during normal phonological processing (e.g., Devlin et al., 2003; McDermott et al., 2003; Seghier et al., 2004). The TMS studies were needed to ascertain whether and where regions from fMRI studies contribute to phonological performance (e.g., Gough et al., 2005; Sliwinska et al., 2015). Based on these results, the current study has shown how lesion analyses with stroke patients identified the extent of grey/white matter that is required for phonological processing and how these extended regions were more predictive of behaviour when they were informed by the TMS sites rather than the fMRI sites. For example, in the current study, the TMS-guided pOp region extended medially into dorsal (e.g., arcuate fasciculus) and ventral (e.g., uncinated fasciculus) white matter tracts, whereas the fMRI-guided pOp region extended more laterally and anteriorly into the ventral sensorimotor cortex (while sparing ventral white matter tracts).

4.5.2 The depth and extent of the TMS effect on phonological processing

The finding that the critical lesion sites included the white matter underlying SMG and pOp raises the possibility that the effect of TMS on phonological processing in

neurologically-normal individuals might also be emerging from disruption to much more extensive areas than the recorded site of stimulation. Indeed, contrary to earlier studies where the focality of the electromagnetic fields induced by the TMS coil was estimated to range between 5-10 mm (Brasil-Neto et al., 1992a, b; Wilson et al., 1993; Ravazzani et al., 1996; Thielscher and Kammer, 2002), more recent research into the biophysical aspects of TMS has suggested that in addition to the local effects of magnetic pulses on the targeted brain region, more remote ones could also take place via spreading of TMS-evoked activity to distal but highly interconnected (both structurally and functionally) cortical areas that form part of the same network (Siebner et al., 2009; Wagner et al., 2009; Liew et al., 2014; Nummenmaa et al., 2014; Neggers et al., 2015). This assertion, however, needs to be qualified by the fact that in non-human primates the direct propagation of electrical-stimulation-induced-activity is typically restricted to the next synapse (i.e. to monosynaptic connections) due to synaptic inhibition by GABAergic interneurons (unless stimulation is of very high frequency; i.e. > 200 Hz) (Logothetis et al., 2010). In the context of SMG and pOp specifically, it is therefore plausible to hypothesise that spread of electrical-stimulation-induced-activity may potentially occur based on of the following four lines of evidence: (i) Petrides & Pandya (2009) have shown in non-human primates that SMG and pOp are connected via long monosynaptic association tracts (i.e. through the superior longitudinal fasciculus and the arcuate fasciculus); (ii) Martino et al. (2013) have confirmed the existence of similar connections (i.e. superior longitudinal fasciculus and arcuate fasciculus) between SMG and pOp in humans; (iii) Matsumoto et al. (2004) have demonstrated that electrical stimulation applied directly over pOp in patients with intractable epilepsy elicits neural responses in SMG and vice versa; and, finally, (iv) Hartwigsen et al. (2016) have shown that concurrently targeting SMG and pOp with TMS does not have an additive effect on phonological processing given that, for example, TMS over SMG leads to decreased fMRI activity in pOp (Hartwigsen et al., 2017). Alternatively (or in addition), it could be the case that the reactivity of the whole of the network is altered when one of its nodes is perturbed (e.g., Binney and Lambon Ralph, 2015). In relation to the results of

Experiment II, thus, an impairment could either be due to disruption of the cortex, underlying white matter or both. Future lesion studies are therefore required to compare the phonological processing abilities of stroke patients with focal damage to the grey matter only, white matter only or both. However, this might not be feasible at all since stroke lesions only very rarely affect the grey matter in the absence of co-occurring white matter damage and vice versa.

4.5.3 Performance was not worse following damage to both regions than SMG or pOp alone

In line with the study by Hartwigsen et al. (2016), the findings obtained from Sample 1 (1-5 years post-stroke) indicate that concurrent damage to TMS-guided SMG and pOp did not have an additive effect on the loss of accuracy with which phonologically demanding tasks could be performed. This is not surprising given how consistently damage to each region alone impairs phonological processing abilities during the first 5 years after stroke onset. However, it raises the possibility that concurrent damage to the TMS-guided SMG and pOp regions - both of which extend deep into the underlying white matter - does not have an additive effect on performance because lesions to any part of the dorsal stream (that sever the underlying white matter tracts) impede the cycling of information between frontal and parietal sites. Thus, damage to either region may arguably be sufficient to emulate the disruptive effect of concurrent damage to both SMG and pOp. Likewise, as pointed out by Hartwigsen et al. (2016), TMS of SMG or pOp may also affect the normal functioning of the other region by knocking-out the white matter tracts that connect them.

4.5.4 Focusing on the consistency of lesion effects rather than their statistical significance

The TMS-guided lesion-deficit mapping approach presented here also has advantages relative to other types of analysis techniques whose primary focus is on identifying the most significant across-subject association between lesion and deficit in a purely statistical sense. For example, I found that damage to either TMS-guided SMG

or TMS-guided pOp is sufficient to impair performance on phonologically demanding tasks. Moreover, by deriving a set of measures of classification/prediction accuracy, I have been able to show how damage to either of these regions was very consistently associated with persistent phonological processing impairments during the first 5 years post-stroke (approximately 90% with TMS-guided SMG damage and 80% with TMS-guided pOp damage), which provides a much more appealing explanation of the incidence of aphasic symptomatology than that expected on the basis of recent reviews (Watila and Balarabe, 2015). In addition, by cross-validating the regions in an independent sample of patients who were more than 5 years post-stroke (i.e. Sample 3), I have demonstrated that the consistency of the TMS-guided SMG effect on phonological processing was not affected by time post-stroke (i.e. positive predictive value was still greater than 85%), whereas that of the TMS-guided pOp effect experienced a noticeable drop; presumably because a subset of the pOp patients were able to regain some of their phonological processing abilities over time. Therefore, this novel TMS-guided methodology that focuses on classification/prediction accuracy (expressed as the consistency of the effect across individual patients) rather than group-level statistics (e.g., *t, F* and/or *p* values) may, in the future, help us improve our ability to predict outcome after stroke.

4.5.5 Inter-patient variability in the effect of damage to the TMS-guided regions

Damage to TMS-guided SMG and pOp as defined in this study was associated with impairments on tasks that required phonological processing. However, not all the patients with damage to the TMS-guided regions performed poorly on phonologically demanding tasks. Plausibly, these patients may have had milder impairments that could only be detected by using more sensitive measures such as reaction times or by testing them earlier post-stroke (before recovery had occurred). Nevertheless, the fact that some individuals were able to produce accurate responses (even if they were slower than normal) might reflect inter-subject variability in (i) the functional anatomy of phonological processing (e.g., Price and Friston, 2002; Noppeney et al., 2006; Seghier et al., 2012;

Blank et al., 2017); (ii) the cognitive strategy used for task performance prior to the stroke (e.g., Seghier et al., 2008b; Kherif et al., 2009; Woollams et al., 2017); and/or (iii) the ability to recover phonological processing skills after damage to the TMS-guided regions (e.g., Specht et al., 2009; Abel et al., 2015; Griffis et al., 2017a; Skipper-Kallal et al., 2017a, b). In relation to this last point, it is possible that at least some of the patients might have learnt to recruit an alternative set of areas (e.g., in the right hemisphere) to aid language recovery after a left-hemisphere stroke (e.g., Brownsett et al., 2014; Geranmayeh et al., 2016, 2017; Xing et al., 2016; Lukic et al., 2017). Indeed, it has already been shown by means of TMS that regions within the right SMG and right pOp support normal phonological processing (i.e. in the absence of left-hemisphere damage) (Hartwigsen et al., 2010a, b). Future studies will need to identify the lesion and non-lesion factors that might explain who will and will not be able to regain their phonological processing abilities after full or partial damage to TMS-guided SMG and/or pOp.

4.5.6 Limitations

A potential weakness of the current study is that there was a noticeable fall in the accuracy with which damage to the TMS-guided pOp region predicted the presence of phonological processing impairments when using Sample 3 (> 5 years post-stroke) relative to using Sample 1 (1-5 years post-stroke). This could be an indication of overfitting, which would mean that the TMS-guided pOp region was not only capturing the signal but also the noise in the data from Sample 1, thereby limiting its generalizability (Arbabshirani et al., 2017). However, such an explanation would be at odds with three other observations: (i) damage to TMS-guided SMG remained a very reliable predictor of phonological performance even when tested in Sample 3 and even though TMS-guided SMG was identified using the exact same procedure as for TMS-guided pOp; (ii) analysis of Sample 3 indicated that the TMS-guided SMG and pOp regions were able to explain phonological impairments in 94% of the affected patients, and (iii) the ability to discriminate the presence or absence of phonological impairments was better when

using the TMS-guided regions than the original TMS regions or fMRI-guided regions, irrespective of time post-stroke (i.e. in Samples 1 and 3; see Tables 4.7 and 4.9).

A more plausible account of the loss of prediction accuracy in the case of the TMS-guided pOp region would be that a subset of the patients with pOp damage were able to regain their phonological processing skills after years of recovery time. Indeed, preliminary results (reported here) from an ongoing longitudinal study appear to support the idea that recovery of language function could explain, at least in part, why the consistency of the TMS-guided pOp effect decreases over time. Furthermore, recent evidence suggests that, contrary to what is commonly assumed, patients continue to recover their language abilities even years after stroke onset (Holland et al., 2017; Hope et al., 2017). Such long-term recovery of language function has been shown to be contingent upon the successful recruitment of intact regions in the damaged hemisphere (typically the left) as well as those in the undamaged hemisphere (typically the right) (e.g., Brownsett et al., 2014; Geranmayeh et al., 2016, 2017; Xing et al., 2016; Lukic et al., 2017). In line with these findings, previous TMS studies of neurologically-normal individuals have associated regions within the right SMG and right pOp with phonological processing (Hartwigsen et al., 2010a, b) and shown that the influence of right on left pOp activity is associated with more efficient performance on phonologically demanding tasks (Hartwigsen et al., 2013). Future studies will need to investigate (i) the lesion and non-lesion factors that might indicate who will and will not recover from damage to TMS-guided pOp; (ii) the brain regions that can support phonological processing following TMS-guided pOp lesions; and (iii) if and why less patients recover from damage to the TMS-guided SMG region than the TMS-guided pOp region.

Another potential limitation is that not all the patients who were impaired on the digit span and non-word reading tasks had above-threshold damage to TMS-guided SMG or TMS-guided pOp. In these patients, it is therefore necessary to determine whether the aphasic symptoms were caused by (i) damage to other regions that form part of the phonological system but that were not the focus of the current study; (ii)

inconsistent effects of partial damage to TMS-guided SMG and pOp; and/or (iii) a combination of damage to areas that support visual and auditory perception and/or speech production.

Finally, the current study does not discount the possibility that identical pOp and SMG regions could have been identified utilising voxel-based lesion-deficit mapping approaches (e.g., voxel-based lesion-symptom mapping or voxel-based morphometry) or that the same classification accuracy could have been obtained by using regions derived from other lesion-deficit mapping analyses such as an “unguided” lesion overlap map of patients with phonological impairments. These issues will be covered in the next chapter (i.e. Experiment III).

4.6 Conclusions

This study investigated whether the classification/prediction accuracy of the incidence of phonological processing impairments could be further improved by adapting the borders of the original TMS SMG and pOp sites (from Experiment I) to include the surrounding grey and white matter. It was shown that the classification accuracy of these new “TMS-guided” regions (which extended deep into the underlying white matter) was more than 3 times higher than that of the original TMS regions irrespective of time since stroke. Likewise, the ability to discriminate between those with and without phonological impairments was substantially better when using TMS-guided SMG and pOp regions than fMRI-guided SMG and pOp regions. In addition, it was found that the mapping between TMS-guided pOp lesions and phonological impairments was less consistent in patients who were more than 5 years versus 1 to 5 years post-stroke. Preliminary longitudinal evidence suggested that this probably occurred because a subset of patients with TMS-guided pOp damage were able to regain their phonological processing skills after years of recovery time.

In summary, the novel contribution of the current study is to identify the extent of pOp and SMG damage that impairs accurate phonological processing, even years after stroke. Collectively, the results appear to indicate (i) that the critical lesion sites involve

a combination of grey and white matter and (ii) that other brain areas are not typically able to consistently compensate for the contribution that these TMS-guided regions make to phonological processing. These findings may therefore have important clinical implications for predicting the incidence of phonological processing impairments in future patients.

CHAPTER 5 (Experiment III):

Comparing the Prediction Accuracy of the TMS-guided Regions with that of a Set of Regions Derived from Voxel-based Lesion-Deficit Mapping Analyses

5.1 Summary

Experiment II showed that the incidence of phonological processing impairments was predicted better when the SMG and pOp regions were adapted to include the grey and white matter that was consistently damaged in patients with (i) phonological processing impairments and (ii) lesions to one or the other of the spherical TMS regions derived from previous studies of neurologically-normal individuals. Experiment III compared the classification accuracy of the TMS-guided regions to that of three sets of regions obtained from voxel-based lesion-deficit mapping analyses: ROI-1 was derived from an "unguided" lesion overlap map of all patients with selective phonological processing impairments. The area of maximum lesion overlap included both the dorsal and ventral white matter and had less than 20% overlap with either the TMS-guided SMG or pOp regions. ROI-2 comprised regions where the frequency of damage was most significantly different in a voxel-based comparison of lesion sites in two groups of patients - with and without phonological impairments. This lesion analysis identified a region in the mid portion of the left superior temporal gyrus extending into the ventral white matter. ROI-3 involved regions where the degree of structural damage was most significantly correlated with the severity of the functional impairment of interest in a voxel-based multiple regression analysis across patients with and without phonological impairments. The results included 92% of the voxels in ROI-2 (i.e. ROI-2 and ROI-3 were very similar).

Comparison of the prediction accuracy for the new regions with that for the TMS-guided regions showed that the presence/absence of phonological impairments was better predicted by the presence/absence of damage to the TMS-guided regions than any of the other three sets of regions in both Sample 1 (used to define the regions) and Sample 3 (that was entirely independent of the region identification process).

The finding that the TMS-guided SMG and pOp regions improved classification accuracy compared to the unguided-lesion-overlap-map region (ROI-1) illustrates how TMS studies of the healthy brain can be used to guide the identification of regions where brain damage is likely to cause persistent behavioural effects. Critically, however, the

highest classification accuracy was observed when the TMS-guided regions were combined with ROI-2 or ROI-3; because the most significant effects from the voxel-based statistical analyses identified the mid part of the left superior temporal gyrus rather than the pOp and SMG regions of interest from TMS studies of neurologically-normal participants. Consequently, lesions to ROI-2 (or ROI-3) accounted for a unique and significant proportion of the variability in phonological outcomes. This highlights how different approaches to lesion-deficit mapping (including those introduced in Chapter 3 and 4) may help to pinpoint all the critical nodes of a distributed neural system where damage disrupts behaviour.

5.2 Introduction

The goal of my third study was to investigate whether the ability to discriminate between patients with and without phonological impairments was higher when using the TMS-guided regions identified in Experiment II than regions obtained from three different types of lesion analysis that did not depend on previous studies of neurologically-normal individuals. These included: (i) an unguided lesion overlap map from patients with selective phonological impairments; (ii) a voxel-based comparison of lesioned versus non-lesioned voxels in patients with and without phonological impairments; and (iii) a voxel-based multiple regression analysis that revealed where the degree of structural damage correlated with the severity of phonological impairments.

I hypothesised that the classification accuracy of the TMS-guided regions would be higher than that of the alternative set of regions because there is already prior evidence that the TMS regions contribute to normal phonological processing. Alternatively, however, it might be possible to identify the critical lesion sites from highly controlled statistical analyses without the need for regions of interest from TMS studies of neurologically-normal subjects.

The first lesion-deficit mapping analysis involved (a) creating a lesion overlap map from patients with selective phonological processing impairments and (b) extracting the brain areas (if any) that were commonly compromised in these patients. This is a widely used approach that pinpoints the parts of the brain where damage is frequently observed in patients with a deficit of interest. For instance, in a seminal study, Dronkers (1996) overlapped the binary lesion images (manually extracted from CT and MRI scans) of a group of 25 stroke patients who had the same behavioural deficit (i.e. apraxia of speech) and found that a region of the left precentral gyrus of the insula was damaged in all cases (i.e. 100% lesion overlap). The well-recognised problem with such lesion overlap approach is that a lesion overlap map (LOM) from stroke patients inevitably conflates critical regions with non-critical regions that are co-incidentally damaged by stroke but irrelevant for the function being investigated (Kimberg et al., 2007; Inoue et

al., 2014; Mah et al., 2014; Sperber and Karnath, 2017). This is particularly problematic when the same deficit can be caused by damage to two or more different regions (that are each part of a distributed neural system that supports the function of interest) because, when a lesion overlap map is created for all patients with the deficit, the area of maximum lesion overlap may not include the critical lesion sites at all. Instead, the maximum overlap will be skewed towards regions that are most susceptible to stroke damage (Kimberg et al., 2007; Inoue et al., 2014; Mah et al., 2014; Sperber and Karnath, 2017). Specifically, the white matter in the vicinity of the central portion of the middle cerebral artery is particularly vulnerable to stroke damage (e.g., Phan et al., 2005; Stoeckel et al., 2007; Cheng et al., 2011) and might be consistently compromised in patients whose phonological processing impairments are caused by damage to SMG or pOp or both. The TMS-guided regions help to overcome this problem by (i) grouping patients according to damage to "distant" regions of interest, and (ii) removing patients with very large lesions that span multiple sites.

The second lesion-deficit mapping analysis entailed identifying the brain regions where the frequency of damage was greater in patients with selective phonological impairments relative to patients without phonological impairments. Such methodology involves the use of formal statistical machinery to draw inferences about the differential distribution of damage in two groups of tightly matched subjects who are assumed to differ along one dimension only (i.e. usually the presence or absence of a deficit of interest) (Karnath et al. 2004; Rudrauf et al., 2008). For example, Karnath et al. (2004) conducted a voxel-wise statistical analysis that compared the lesion images of a group of 140 right-hemisphere stroke patients, who either did or did not exhibit the symptoms of spatial neglect. It was shown that, relative to the control group (i.e. patients that did not have spatial neglect), those with spatial neglect had a higher frequency of damage to a set of regions encompassing the right superior temporal gyrus, planum temporale, sensorimotor cortex, insula, putamen and caudate nuclei (even after Bonferroni correction for multiple comparisons). Despite the rational logic of this analysis, there are

two reasons why I hypothesised that regions based on univariate voxel-based statistical comparisons of two groups would be less predictive than the TMS-guided regions. The first is the same as that raised above for the unguided lesion overlap map: the spatial extent and location of the identified regions may be distorted when multiple lesion sites (e.g., SMG and pOp) can result in the same deficit. The second is that the inclusion of patients who may have recovered from initial deficits could hamper the identification of necessary regions because these patients would mistakenly be assigned to the control group (Karnath and Rennig, 2017). I, therefore, hypothesised that the regions identified by the voxel-based statistical comparison of two groups (with and without the deficit) would (i) not necessarily be the same as those identified by the TMS-guided approach and (ii) have lower prediction accuracy than the TMS-guided regions.

The third lesion-deficit mapping analysis was a univariate voxel-based correlational methodology that identified the brain regions where the degree of damage to each voxel correlated with the severity of phonological impairments after factoring out other sources of variance (e.g., semantic processing abilities). This type of voxel-based lesion-deficit technique differs from the group comparison described above because patients do not need to be split into impaired and non-impaired subsets. For instance, Bates et al. (2003) found that poor speech fluency was associated with damage to portions of the insula and superior longitudinal fasciculus. Furthermore, by identifying an alternative set of areas where damage was associated with poorer auditory comprehension, the authors were able to confirm that language production and comprehension dissociate at the neural level. More recent developments in univariate lesion-deficit mapping permit (i) the addition of multiple regressors to factor out variance unrelated to the cognitive process being investigated, such as lesion size and behavioural covariates of no interest (e.g., Geva et al., 2011; Schwartz et al., 2012; Pillay et al., 2014) and (ii) the use of continuous measures of structural abnormality (Tyler et al., 2005; Leff et al., 2009; Butler et al., 2014; Gajardo-Vidal, Lorca-Puls et al., 2018; Lorca-Puls et al., 2018). The multiple regression approach was therefore expected to

identify similar regions of interest to the group comparison analysis described above. However, it might be more sensitive to effects of interest than the group comparison because the behaviour of interest (i.e. phonological processing abilities) and structural damage are both measured on a continuous scale (i.e. the severity of impairments and the degree of structural abnormality) rather than a categorical scale (i.e. the presence or absence of an impairment/damage). Conversely, the inclusion of regressors of no interest (e.g., semantic processing abilities) may sensitise the analysis to the effects of interest but might also remove variance of interest.

The decision to perform univariate voxel-based rather than multivariate machine-learning-based lesion-deficit mapping analyses was primarily motivated by the following three observations: (i) multivariate lesion-deficit mapping methods are still in their early days as reflected by the lack of consensus as to how these types of analyses should be conducted (Karnath et al., 2018) and ongoing discussions on how to optimally estimate the statistical significance of the ensuing spatial maps (e.g., Yourganov et al., 2018); (ii) it is not clear how the weights in multivariate models should be interpreted in neurobiological terms since large weights may bear no relationship at all to the cognitive process under study (and vice versa; Haufe et al., 2014) or how the lesion information should be encoded to produce stable results (Rondina et al., 2016); and, finally, (iii) as mass-univariate analyses, popular multivariate implementations are prone to displacement of lesion-deficit mappings towards areas of greater susceptibility to vascular events (Sperber et al., 2018) and can fail to capture a substantial proportion of the variance they are designed to explain (Zhang et al., 2014; Pustina et al., 2018). Hence, although multivariate machine-learning-based lesion-deficit mapping techniques hold great promise and may eventually replace their univariate counterparts, the current state of affairs makes it difficult to exploit their full potential. Moreover, the overarching goal of the current chapter was to evaluate the added value of the TMS-guided lesion-deficit mapping approach introduced in Experiment II against a representative collection of techniques. This condition seems at present to be better satisfied by univariate voxel-

based rather than multivariate machine-learning-based methods. In addition and considering the distributed nature of the brain systems underpinning human cognitive functions, I formally assessed the possibility that prediction accuracy may be highest when the joint (or multivariate) contribution of multiple regions of interest is taken into account, which is somewhat equivalent to carrying out a network analysis.

In summary, each of these approaches to lesion-deficit mapping have attracted some level of criticism for one reason or another (for more details, see Bates et al., 2003; Hillis et al., 2004; Inoue et al., 2014; Mah et al., 2014; Price et al., 2017; Gajardo-Vidal, Lorca-Puls et al., 2018; Lorca-Puls et al., 2018). The critical point, however, is that univariate lesion-deficit mapping analyses are still widely used by and remain very popular among researchers (de Haan and Karnath, 2018), which makes them a representative collection of techniques. Furthermore, the focus of the current study was not to determine the best way to assess lesion-deficit associations but to compare the classification accuracy of the TMS-guided regions with that of an alternative set of regions derived from a variety of voxel-based lesion-deficit mapping analyses. In this sense, previous TMS and fMRI studies of neurologically-normal subjects have implicated multiple brain areas in phonological processing (e.g., Devlin et al., 2003; McDermott et al., 2003; Acheson et al., 2011; Krieger-Redwood et al., 2013; Murakami et al., 2015; Hartwigsen et al., 2016). From a neural network perspective, it is therefore possible to anticipate that if different lesion-deficit mapping methodologies identify different critical nodes of the phonological system where brain damage impairs behaviour, then the highest classification accuracy should be observed when the regions are considered in combination rather than in isolation. Conversely, if highly overlapping regions are detected or if a subset of the regions consists of areas that are commonly compromised by stroke but do not include the full extent of damage that is necessary for the presence of persistent phonological impairments, then the accuracy of the classification/prediction will not necessarily benefit from using the regions in combination because they are redundant or have limited predictive power.

5.2.1 Research question

In short, the current experiment attempted to address the following research question:

- How does the classification accuracy of the TMS-guided regions compare to that of an alternative set of regions derived from voxel-based lesion-deficit mapping analyses?

5.3 Materials and Methods

5.3.1 Participants

Samples 1 (1-5 years post-stroke) and 3 (> 5 years post-stroke) in the current chapter were the same as Samples 1 and 3 in the previous chapter (i.e. Experiment II). Sample 1 was used to test how consistently damage to distinct regions of interest derived from various types of lesion-deficit mapping analyses was associated with the presence or absence of phonological processing impairments. Sample 3 was used to validate the lesion-deficit associations identified with Sample 1, while also considering the effect of time post-stroke (> 5 years versus 1-5 years). There was no overlap between Samples 1 and 3.

As in Experiments I and II, two subsets of patients were selected from Sample 1 to determine the percentage of damage to any given region that best accounted for the presence or absence of phonological impairments. Subset 1A included those who were categorised with phonological impairments but not semantic impairments (see Chapter 3). Subset 1B included those who were matched to Subset 1A for left-hemisphere lesion size but did not meet the criteria used to define phonological impairments. Lesion size was matched between the two groups by finding the minimum and maximum lesion volumes that were common to both groups with no significant differences in mean lesion size across groups. For more details on Subsets 1A and 1B, see Table 3.4 in Chapter 3.

5.3.2 Regions of interest from univariate lesion-deficit mapping analyses

There are multiple ways in which lesion-deficit analyses can be conducted including lesion overlap maps of patients of interest and distinct types of statistical comparisons (group or correlational) of patients with and without deficits. These analyses can also be conducted with different types of behavioural data (presence/absence of deficit or severity of deficit) and different types of lesion images (presence/absence of damage or degree of damage) (for more information, see Rorden et al., 2007; de Haan and Karnath, 2018; Karnath et al., 2018; Sperber and Karnath, 2018).

I chose to draw on three clearly distinguishable methodologies that are widely used by the lesion-deficit mapping community, thereby providing a representative collection of techniques. The classification accuracy for each set of regions was compared to that of the TMS-guided regions (i.e. those identified in Experiment II). Additionally, I explored the possibility that using a combination of regions may provide a better fit of the data than considering each of the different types of regions alone.

The first region of interest (ROI-1) was derived from an overlap map of the lesion images from all 23 patients in Subset 1A (categorised with phonological impairments but not semantic impairments); see Figure 5.1A. The purpose of this "unguided" lesion overlap map analysis was to identify the brain regions most commonly damaged in patients with selective phonological impairments. It was found that the maximum number of patients who had damage at any given voxel was 19 (out of 23). This degree of overlap, however, was only observed in a very small region (i.e. 0.1 cm^3). Comparison of the classification accuracy for different degrees of overlap identified that an overlap of 16 patients (out of 23), observed over a 7.3 cm^3 region (see ROI-1 in Figure 5.1B), was the best predictor of phonological performance.

The second region of interest (ROI-2) was identified using a univariate voxel-based group comparison of the frequency of lesions (as encoded by the binary lesion images) in the 23 patients with phonological but not semantic impairments (Subset 1A) and the 32 patients who did not meet the criteria for phonological impairments (Subset

1B). The analysis (i) was conducted in NPM (non-parametric mapping) which is part of the MRIcron software package (https://www.nitrc.org/projects/mricron); and (ii) followed procedures described in Rorden et al. (2007) including (a) using the Liebermeister test (which has been shown to be more sensitive than the Chi-Squared or Fisher's Exact test) and (b) limiting the analysis to voxels that were damaged in at least 20% of the 55 patients (for rationale, see Sperber and Karnath, 2017). The resulting region of interest (1.0 cm^3) comprised all the voxels that surpassed a statistical threshold of $p < 0.001$ uncorrected, one-tailed (see ROI-2 in Figure 5.1C); because no effects survived a voxel-level family-wise error (FWE) corrected threshold (i.e. $p < 0.05$, FWE-corrected).

The third region of interest (ROI-3) was derived from a univariate voxel-based multiple regression analysis that identified voxels where greater lesion load was significantly associated with poorer phonological processing abilities across the 55 patients in Subsets 1A and 1B (combined). This was performed in SPM12 using multiple regression and the general linear model. The imaging data entered into the analysis were the continuous (fuzzy) lesion images that index the degree of structural abnormality on a continuous scale from 0 (completely normal) to 1 (completely abnormal) at each and every voxel of the brain relative to normative data drawn from a sample of 64 neurologically-normal controls (for more details, see Chapter 2). The continuous lesion images provided a richer level of information than the binary lesion images used in the group comparison described above (but note that either type of image could be used in either analysis). The behavioural regressor of interest was the average scores of non-word reading and digit span, which are sensitive to phonological processing impairments. In addition, the following regressors were included to factor out other sources of variance: (i) auditory word-to-picture matching scores which are sensitive to auditory recognition of aurally presented words and lexical-semantic processing, (ii) visual word-to-picture matching scores which are sensitive to visual recognition of written words and lexical-semantic processing and (iii) visual semantic associations scores, which are sensitive to picture recognition and semantic processing. As in the voxel-based group comparison

described above, the search volume for the voxel-based correlational analysis was restricted to voxels that were damaged in at least 20% of the 55 patients from Subsets 1A and 1B (for rationale, see Sperber and Karnath, 2017). The statistical threshold was set at $p < 0.001$ uncorrected, one-tailed (see ROI-3 in Figure 5.1D); because no significant effects were observed at a voxel-level FWE-corrected threshold of $p < 0.05$.

5.3.3 Determining the threshold for critical damage

For each region, I investigated how accurately different thresholds of damage (e.g., 100%, 90%, 80% and 70%) categorised 55 patients in Subsets 1A and 1B into those with versus without phonological impairments; see Chapters 3 for details.

5.3.4 Evaluating the combination of regions that maximised classification accuracy

A logistic regression framework was used to formally evaluate whether the presence or absence of above-threshold damage to each region made a unique and significant contribution to classification accuracy. Specifically, the binary outcome variable was the presence or absence of phonological impairments (as defined by abnormally low performance on both the non-word reading and digit span tasks), and the predictors were the presence or absence of above-threshold damage to each set of regions. This analysis was conducted on Sample 1 (1-5 years post-stroke) and Sample 3 (> 5 years post-stroke) separately.

5.4 Results

5.4.1 Regions of interest from univariate lesion-deficit mapping analyses

ROI-1 (the unguided-LOM region) included portions of the white matter between the TMS-guided SMG and pOp regions in the anterior, posterior and long segments of the arcuate fasciculus, the corticospinal tract, inferior longitudinal fasciculus, inferior fronto-occipital fasciculus and periventricular white matter (see Figure 5.1B). Critically, less than 20% of either TMS-guided SMG or TMS-guided pOp was contained within ROI-1 (see Table 5.1). In contrast, greater damage to ROI-1 co-occurred with greater damage to the TMS-guided regions (see Figure 5.2).

ROI-2 (from the voxel-based group comparison of patients with versus without phonological impairments) was located in the mid part of the left superior temporal gyrus including portions of the planum temporale and Heschl's gyrus as well as portions of the inferior longitudinal fasciculus, inferior fronto-occipital fasciculus and hippocampus (see Figure 5.1C). This region is completely different from ROI-1 and the TMS-guided regions (see Table 5.1). Importantly, the degree of damage to ROI-2 co-varied with that to the TMS-guided regions (see Figure 5.2), and significantly more so for the SMG region than the pOp region ($z = 4.42$, $p < 0.001$ in Sample 1; $z = 2.83$, $p = 0.005$ in Sample 3).

ROI-3 (where greater lesion load was associated with poorer phonological processing abilities) was similar to ROI-2 in that it was centred on the mid portion of the left superior temporal gyrus (see Figure 5.1D). Specifically, 92% of ROI-3 was contained within ROI-2 (see Table 5.1). As with ROI-2, greater lesion load in ROI-3 was more strongly correlated with greater lesion load in the TMS-guided SMG region than the TMS-guided pOp region ($z = 4.20$, $p < 0.001$ in Sample 1; $z = 3.54$, $p < 0.001$ in Sample 3); see Figure 5.2.

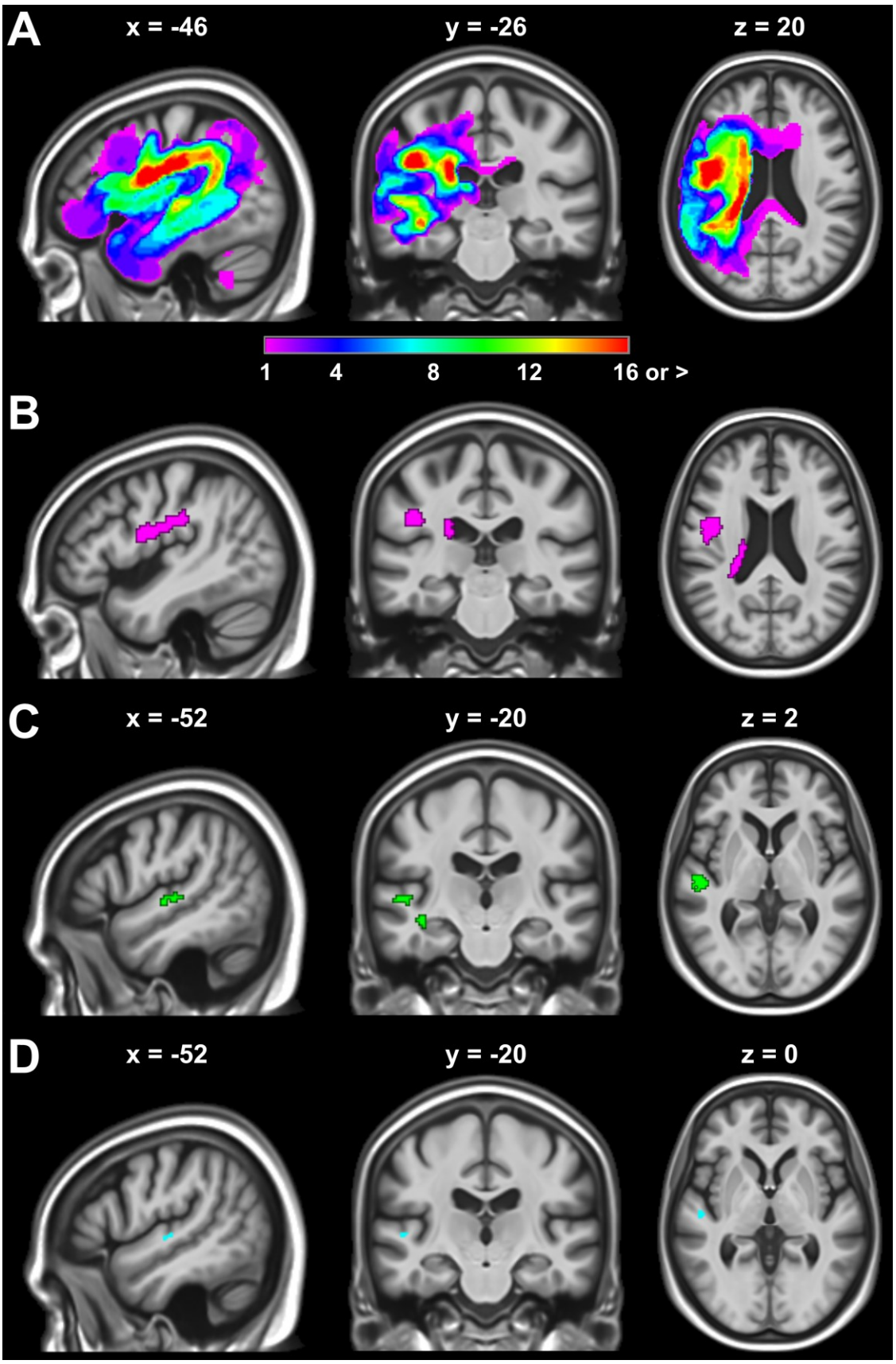

Figure 5.1: Univariate lesion-deficit mapping regions. (A) Lesion overlap map of patients with selective phonological processing impairments in Subset 1A (*n* = 23). (B) ROI-1 (in violet) only included voxels that were damaged in 16 out of 23 patients from the lesion overlap map. The voxel-based lesion-deficit mapping regions are shown in green for ROI-2 (C) and cyan for ROI-3 (D).

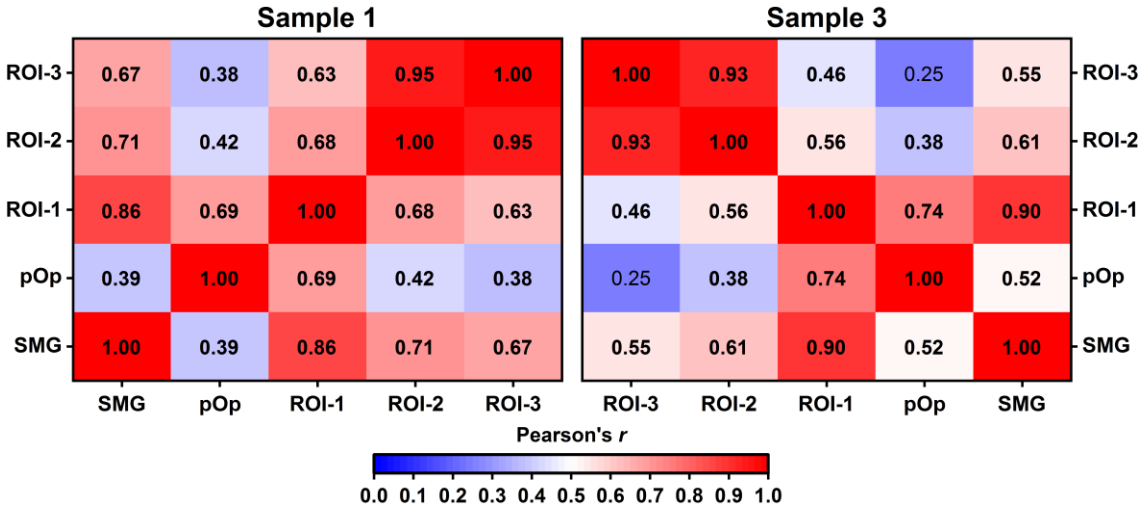

Figure 5.2: Co-occurring patterns of damage. The figure illustrates how likely it was that any given pair of regions were concurrently affected by stroke damage in Samples 1 and 3. All regional lesion load correlations were statistically significant at $p < 0.001$, except for that between TMS-guided pOp and ROI-3 in Sample 3 (i.e. $r(108) = 0.25$, $p = 0.008$).

Table 5.1: Overlap between TMS-guided and univariate lesion-deficit mapping regions.

		TMS-guided		Univariate		
		SMG	**pOp**	**ROI-1**	**ROI-2**	**ROI-3**
TMS-guided	**SMG**	100%	0%	61%	4%	0%
	pOp	0%	100%	24%	0%	0%
Univariate	**ROI-1**	18%	8%	100%	2%	8%
	ROI-2	0%	0%	0%	100%	92%
	ROI-3	0%	0%	0%	9%	100%

The table shows the percentage of each of the regions listed at the top contained within each of the regions on the left-hand side.

5.4.2 Critical damage thresholds

The thresholds for the degree of damage that best explained the presence or absence of phonological impairments was: 90%, 70% and 100%, for ROI-1, ROI-2 and ROI-3, respectively (see Table 5.2). Henceforth, “above-threshold damage” refers to instances where the degree of damage to the region of interest was equal to or greater than the region-specific threshold.

Table 5.2: Critical damage thresholds for univariate lesion-deficit mapping regions.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
ROI-1	100%	100%	63%	17%	100%	---
	90%	70%	64%	30%	91%	***4.2***
	80%	56%	65%	43%	75%	2.3
	70%	52%	67%	57%	63%	2.2
ROI-2	100%	75%	61%	13%	97%	4.7
	90%	67%	63%	26%	91%	3.4
	80%	69%	67%	39%	88%	4.5
	70%	71%	68%	43%	88%	***5.4***
ROI-3	***100%***	75%	72%	52%	88%	***7.6***
	90%	72%	73%	57%	84%	7.0
	80%	72%	73%	57%	84%	7.0
	70%	70%	74%	61%	81%	6.7

Threshold = percentage of the region of interest (ROI) damaged; PPV/NPV = positive and negative predictive values.

5.4.3 Classification accuracy for Sample 1

For ROI-1, phonological impairments were observed in 89% (33/37) of the patients with above-threshold damage and 24% (28/117) of the patients with below-threshold damage. For ROI-2, phonological impairments were observed in 85% (33/39) of the patients with above-threshold damage and 24% (28/115) of the patients with below-threshold damage. For ROI-3, phonological impairments were observed in 81% (35/43) of the patients with above-threshold damage and 23% (26/111) of the patients with below-threshold damage (see Table 5.3).

The odds ratios for ROI-1, ROI-2 and ROI-3 were: 26.2, 17.1 and 14.3, all of which were lower than that for the TMS-guided regions (i.e. 32.6); see Table 5.3. However, the odds ratio experienced a marked increase when the TMS-guided regions were combined with ROI-2 (49.2) or ROI-3 (47.5). This is because ROI-2 and ROI-3 accounted for 7 and 8 patients with phonological impairments who did not have above-threshold damage to the TMS-guided regions, respectively. Conversely, the TMS-guided regions accounted for 20 patients with phonological impairments who did not have above-threshold damage to ROI-2, 19 patients who did not have above-threshold damage to ROI-3 and 16 patients who did not have above-threshold damage to ROI-1.

There was no further improvement when ROI-1 was included in any combination of the other ROIs or when ROI-2 and ROI-3 were added to the same analysis (see Table 5.3). In other words, the best classification accuracy was obtained from the combination of the TMS-guided regions with ROI-2 (or ROI-3).

Table 5.3: Classification accuracy for TMS-guided and univariate lesion-deficit mapping regions in Sample 1.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
(1) TMS-guided	80%	85%	85%	75%	91%	32.6
(2) ROI-1	90%	89%	76%	54%	96%	26.2
(3) ROI-2	70%	85%	76%	54%	94%	17.1
(4) ROI-3	100%	81%	77%	57%	91%	14.3
(1) & (2)	as above	83%	87%	80%	89%	33.9
(1) & (3)	as above	83%	91%	87%	88%	49.4
(1) & (4)	as above	81%	92%	89%	86%	47.5
(2) & (3)	as above	85%	80%	66%	92%	23.4
(2) & (4)	as above	82%	82%	69%	90%	20.6
(3) & (4)	as above	82%	77%	59%	91%	15.3
(1) & (2) & (3)	as above	82%	91%	87%	87%	44.7
(1) & (2) & (4)	as above	79%	92%	89%	85%	43.5
(1) & (3) & (4)	as above	81%	92%	89%	86%	47.5
(2) & (3) & (4)	as above	82%	82%	69%	90%	20.6
(1) & (2) & (3) & (4)	as above	79%	92%	89%	85%	43.5

Threshold = critical damage threshold for each region of interest (ROI); PPV/NPV = positive and negative predictive values.

5.4.4 Classification accuracy for Sample 3

Although Sample 3 was entirely independent of the region identification process, the classification accuracy of the TMS-guided regions (odds ratio = 27.6) was still substantially better than that of ROI-1 (odds ratio = 10.8), ROI-2 (odds ratio = 5.7) or ROI-3 (odds ratio = 7.8). The best fit of the data, however, was when the analysis included the TMS-guided regions as well as ROI-3 (odds ratio = 55.7), with no further improvements observed when the TMS-guided regions were combined with ROI-1 (odds ratio = 39.7) or ROI-2 (odds ratio = 30.7) or when any other combination of ROIs was used (see Table 5.4).

The combination of the TMS-guided regions and ROI-3 resulted in very high sensitivity because 98% (39/40) of patients with phonological impairments had above-threshold damage to one or more of these regions. Put another way, the absence of above-threshold damage to these regions predicted the absence of phonological impairments in 98% (40/41) of patients (i.e. very high negative predictive value). The positive predictive value, however, was much lower because 42% (28/67) of patients with above-threshold damage to the regions of interest did not have phonological impairments (see Table 5.5 for full breakdown).

Table 5.4: Classification accuracy for TMS-guided and univariate lesion-deficit mapping regions in Sample 3.

ROI	Threshold	PPV	NPV	Sensitivity	Specificity	Odds Ratio
(1) TMS-guided	80%	64%	94%	93%	69%	27.6
(2) ROI-1	90%	72%	81%	65%	85%	10.8
(3) ROI-2	70%	66%	75%	53%	84%	5.7
(4) ROI-3	100%	69%	78%	60%	84%	7.8
(1) & (2)	as above	63%	96%	95%	68%	39.7
(1) & (3)	as above	59%	95%	95%	62%	30.7
(1) & (4)	as above	58%	98%	98%	59%	55.7
(2) & (3)	as above	64%	84%	75%	75%	9.0
(2) & (4)	as above	65%	86%	80%	75%	12.0
(3) & (4)	as above	65%	77%	60%	81%	6.3
(1) & (2) & (3)	as above	59%	95%	95%	62%	30.7
(1) & (2) & (4)	as above	58%	98%	98%	59%	55.7
(1) & (3) & (4)	as above	58%	98%	98%	59%	55.7
(2) & (3) & (4)	as above	63%	86%	80%	72%	10.3
(1) & (2) & (3) & (4)	as above	58%	98%	98%	59%	55.7

Threshold = critical damage threshold for each region of interest (ROI); PPV/NPV = positive and negative predictive values.

Table 5.5: Incidence of phonological processing impairments according to the damage status of the best combination of regions.

Above-threshold damage to:	Phonological Impairments?	
	Yes	**No**
Sample 1 (1-5 years post-stroke)		
(1) TMS-guided SMG only	5	0
(2) TMS-guided pOp only	13	5
(3) ROI-2 only	7	3
(1) & (2) not (3)	2	0
(1) & (3) not (2)	4	1
(2) & (3) not (1)	7	0
(1) & (2) & (3)	15	2
Neither (1), (2) or (3)	8	82
Error	13%	12%
Sample 3 (> 5 years post-stroke)		
(1) TMS-guided SMG only	1	1
(2) TMS-guided pOp only	9	12
(3) ROI-3 only	2	7
(1) & (2) not (3)	5	4
(1) & (3) not (2)	7	0
(2) & (3) not (1)	4	1
(1) & (2) & (3)	11	3
Neither (1), (2) or (3)	1	40
Error	3%	41%

The error term represents the percentage of those patients with phonological impairments who did not have above-threshold damage to any of the regions (i.e. false negative rate; left-hand side column) as well as the percentage of those patients without phonological impairments who had above-threshold damage to one or more of the regions (i.e. false positive rate; right-hand side column).

5.4.5 Logistic regression analysis

In Sample 1, the full logistic regression model predicted with an overall accuracy of 88% the presence or absence of phonological impairments relative to 60% for an intercept only model ($X^2(3) = 84.27$, $p < 0.001$, Nagelkerke $R^2 = 57.0\%$). Crucially, each of the variables included in the regression equation (i.e. TMS-guided pOp, TMS-guided SMG and ROI-2) made a unique and significant contribution to prediction accuracy, over and above that added by the other regions; and even after controlling for the effect of lesion size ($p < 0.05$ for all).

In Sample 3, the full model yielded a prediction accuracy (i.e. 81%) which was significantly higher than that (i.e. 63%) of an intercept only model ($X^2(3) = 47.26$, $p < 0.001$, Nagelkerke $R^2 = 48.4\%$). More importantly, above-threshold damage to TMS-guided pOp, TMS-guided SMG or ROI-3 was significantly and independently associated with the presence of phonological impairments, even after factoring out variance accounted for by lesion size ($p < 0.05$ for all).

5.5 Discussion

Experiment III compared the predictive power of the TMS-guided regions (from Experiment II) with that of an alternative set of regions derived from a representative collection of voxel-based lesion-deficit mapping analyses. I found that the ability to discriminate patients with versus without phonological impairments was substantially better when using the TMS-guided regions identified in Experiment II than regions obtained from: (i) an unguided lesion overlap map (LOM); (ii) a voxel-based comparison of lesions in patients with versus without phonological impairments; and (iii) a voxel-based multiple regression analysis. Moreover, when the anatomical location of these regions was compared, it was shown that the degree of overlap between any given pair varied greatly, ranging from as much as 92% (e.g., for ROI-2 and ROI-3) to as little as 0% (e.g., for TMS-guided and ROI-2 or ROI-3); see Table 5.1. More importantly, however, the combination of the non-overlapping TMS-guided pOp, TMS-guided SMG and ROI-2/ROI-3 regions yielded the highest classification accuracy, because each region explained a unique and significant proportion of the variance in phonological outcome irrespective of time post-stroke (1-5 years versus > 5 years) and lesion size.

5.5.1 The lesion sites that best explain the incidence of phonological impairments

The presence or absence of phonological impairments was explained by the TMS-guided regions better than any of the other sets of regions (see Tables 5.3 and 5.4). However, the best fit of the data was when the analysis took into account ROI-2 or ROI-3 (centred on the mid portion of the left superior temporal gyrus) as well as the TMS-guided regions (pOp and SMG). This is because damage to the TMS-guided regions

explained phonological impairments in patients who did not have damage to ROI-2 or ROI-3; while, conversely, damage to ROI-2 or ROI-3 explained phonological impairments in patients who did not have damage to the TMS-guided regions (see Table 5.5 for full breakdown). Together, the TMS-guided regions and ROI-2 or ROI-3 were able to account for the incidence of phonological impairments after stroke in most of the affected patients (i.e. sensitivity 87% in Sample 1 and 98% in Sample 3; see Tables 5.3 and 5.4). Put another way, the absence of damage to these regions predicted the absence of phonological impairments very accurately (i.e. negative predictive value 91% in Sample 1 and 98% in Sample 3; see Tables 5.3 and 5.4). This is quite remarkable in that it provides a much more consistent explanation of the incidence of aphasic symptomatology than that expected on the basis of recent reviews (Watila and Balarabe, 2015) and may, in future, help to improve our ability to predict positive outcomes after stroke. Moreover, it highlights how distinct regions of interest derived from a variety of lesion-deficit mapping analyses can each contribute unique information, which when integrated could help to pinpoint all the critical nodes of a distributed neural system where brain damage disrupts behaviour (Gajardo-Vidal, Lorca-Puls et al., 2018).

5.5.2 The seed TMS regions increased classification accuracy

The functionally localised TMS regions based on studies of neurologically-normal individuals (from Gough et al., 2005 and Sliwinska et al., 2015) were essential for stratifying the patients into different groups (with damage to pOp or SMG) before generating lesion overlap maps. Without such guidance, a lesion overlap map identified brain areas that were commonly compromised in patients with selective phonological impairments, but the region of interest derived from the unguided lesion overlap map (i.e. ROI-1) did not include the full extent of pOp and SMG damage that was necessary for the presence of persistent phonological impairments. Instead, the unguided-lesion-overlap-map region (ROI-1) included the white matter tracts running between SMG and pOp. Consequently, damage to ROI-1 was highly likely in patients with SMG or pOp lesions (see Figure 5.2). The involvement of these white matter tracts may merely reflect

greater susceptibility to vascular damage caused by stroke (Phan et al., 2005; Stoeckel et al., 2007; Cheng et al., 2011; Mah et al., 2014; Sperber and Karnath, 2016) because the lesion overlap map approach cannot exclude the possibility that the same white matter areas are as frequently damaged in patients without phonological impairments as they are in patients with phonological impairments. In brief, the inferred spatial extent and location of critical regions in the unguided lesion overlap map may have been distorted towards regions that are commonly compromised by stroke (Kimberg et al., 2007; Inoue et al., 2014; Mah et al., 2014; Sperber and Karnath, 2017).

5.5.3 The voxel-based statistical analyses failed to detect the TMS-guided regions

The univariate voxel-based analyses (used to identify ROI-2 and ROI-3) failed to identify the TMS-guided SMG and pOp regions that have been found to be important for normal phonological processing. This is unlikely to be due to limited statistical power because there were plenty of patients with damage to both SMG and pOp (see Figure 3.8 in Chapter 3). Instead, it suggests an inconsistent mapping of damage/preservation with the presence or absence of phonological impairments. For example, damage to a region may not always be associated with a deficit of interest if the analysis includes patients who have recovered from an earlier deficit (Karnath and Rennig, 2017). In addition, preservation of a region will not be associated with spared performance when the same functional impairment can be caused by more than one lesion site (Gajardo-Vidal, Lorca-Puls et al., 2018) as appears to be the case with TMS-guided SMG and pOp. This is particularly problematic for univariate voxel-based analyses that assume that the effect of damage to one voxel is not influenced by the lesion status of any other voxel in the brain (i.e. thousands of independent statistical tests are conducted on a voxel-by-voxel basis) (DeMarco and Turkeltaub, 2018). Thus, when damage to more than one region (e.g., region A or B) can independently cause the same functional impairment, univariate voxel-based analysis will fail to identify regions A and B (unless the frequency of damage to A & B is substantially greater than that to either region alone).

This is because preservation of region A will inconsistently be associated with preserved function if region B is damaged and vice versa.

The importance of SMG and pOp for phonological processing might also have been missed because the univariate statistical analyses used to identify ROI-2 and ROI-3 were not able to determine how the combination of damage to two different regions impacts upon outcome and recovery (Price et al., 2017). For example, the results of Experiment II and III suggest that it is the combination of damage to cortical regions (SMG or pOp) and the underlying white matter that causes persistent phonological impairments. This is related to the so called “partial injury problem”, whereby a functional impairment is only observed when a certain portion of a critical region encompassing multiple voxels is damaged (Rorden et al., 2009). The problem arises because the association between lesion and deficit appears to be inconsistent if the effect of damage on behaviour is assessed at the level of single voxels. In other words, the ability of univariate voxel-based analyses to detect statistically significant effects is compromised by the fact that partial damage to different parts of TMS-guided SMG or pOp does not consistently result in phonological processing impairments. However, if the contribution of multiple voxels (or patches of voxels) is considered simultaneously, the problem becomes solvable. Future statistical lesion-deficit analyses may therefore need to use multivariate rather than univariate methods to, for instance, test whether the effect of damage to one or more regions is super-additive of the effect of damage to either region alone.

5.5.4 The voxel-based statistical analyses identified similar regions

Whereas less than 10% of ROI-2 or ROI-3 were contained within ROI-1 or the TMS-guided regions, there was a substantial degree of overlap between ROI-2 and ROI-3 (see Table 5.1). Specifically, most of the voxels comprising ROI-3 (i.e. 92%) were also part of ROI-2. This is not surprising given that the voxel-based analyses that identified ROI-2 and ROI-3 were designed to answer essentially the same question: where in the

brain does the presence or absence of damage is associated with the presence or absence of phonological impairments?

There were also some differences between ROI-2 and ROI-3. For instance, ROI-2 covered a larger area than ROI-3 (1.0 cm^3 versus 0.1 cm^3) including the underlying white matter and portions of the hippocampus. As these regions were based on results that were thresholded at $p < 0.001$ uncorrected (given that no voxels survived correction for multiple comparisons), the differences between ROI-2 and ROI-3 might represent false positives. Future studies are therefore required to investigate whether the dissimilarities between ROI-2 and ROI-3 occurred because the group comparison that identified ROI-2 was performed using binary measures of structural abnormality (i.e. binary lesion images), while the multiple regression analysis that identified ROI-3 was conducted on continuous measures of structural abnormality (i.e. fuzzy lesion images). In addition, there were discrepancies in how the analyses controlled for co-occurring semantic impairments: the group comparison only looked at patients with phonological impairments in the context of normal semantic performance. In contrast, the multiple regression analysis factored out additional residual variance related to inter-patient differences in semantic processing abilities with the inclusion of behavioural covariates of no interest.

The important point for the purpose of the current study is, however, that the presence or absence of damage to ROI-2 or ROI-3 provided the same classification of outcomes (with versus without phonological impairments) in 38/44 (86%) and 30/37 (81%) of patients in Samples 1 and 3, respectively. ROI-3 only differed from ROI-2 by (i) explaining three more patients with phonological impairments in Samples 1 and 3 (ROI-2 explained an additional patient in Sample 1) and (ii) misclassifying two more patients without phonological impairments in Samples 1 and 3 (ROI-2 misclassified two additional patients in Sample 3).

5.5.5 Limitations

One potential limitation of the current study is that the odds ratio (i.e. a single metric that reflects the overall performance of a binary predictor/classifier) for each of the four sets of regions suffered a noticeable drop when moving from Sample 1 (1-5 years post-stroke) to Sample 3 (> 5 years post-stroke). This could have been caused by a lack of generalizability of the results from one sample to the next (i.e. overfitting; Arbabshirani et al., 2017). However, the fact that other accuracy metrics exhibited the opposite pattern (see Tables 5.3 and 5.4) seems to go against such an explanation and instead favour an interpretation in terms of an overall effect of time post-stroke (an indirect index of recovery of language function). Put another way, it might be that the number of patients who are able to regain their phonological processing abilities following damage to one or more of the identified regions increases with recovery time. This would be consistent with prior reports that have shown that patients continue to improve their residual language skills even years after stroke onset (Holland et al., 2017; Hope et al., 2017) by, for example, recruiting an alternative set of intact regions (e.g., Brownsett et al., 2014; Geranmayeh et al., 2016, 2017; Xing et al., 2016; Griffis et al., 2017a; Lukic et al., 2017; Skipper-Kallal et al., 2017a, b). Future cross-sectional and longitudinal studies will need to investigate the lesion and non-lesion factors that might explain who will and will not have persistent phonological impairments in those with full or partial damage to the regions that have been found to be critical for accurate phonological processing here.

A second potential limitation is that, although the classification accuracy of the TMS-guided regions was demonstrated to be substantially higher than that of ROI-2 and ROI-3 which were derived from analyses that used the most popular univariate voxel-based lesion-deficit mapping approaches, it remains to be tested whether the results would change if the latest machine learning algorithms were utilised to take into account how the effect of damage to one region depends on that in another by explicitly modelling the multivariate patterns in which brain damage can affect behaviour (Zhang et al., 2014; Hope et al., 2015; Yourganov et al., 2016; Pustina et al., 2018). However, despite their

relative advantages, sophisticated machine learning techniques are not without problems because (i) as in univariate analyses, they depend on operator decisions such as the definition of a single region (be it a single voxel, an anatomically-defined region or a data-defined region); and (ii) the multi-region lesion information extracted can become very complex and non-intuitive because of the high dimensionality that arises when the same deficit can be caused by damage to multiple regions and, conversely, when multiple deficits are associated with damage to the same region. Furthermore, there is no standard way of conducting multivariate lesion analyses since new methods are constantly being developed (e.g., Rondina et al., 2016; Yourganov et al., 2016; Malherbe et al., 2018; Pustina et al., 2018). For the time being, my findings in this study have shown that the presence or absence of damage to the TMS-guided SMG and/or pOp regions is a highly reliable predictor of the incidence of phonological processing impairments.

Finally, the focus of my thesis has so far been on establishing whether TMS (and fMRI) of the healthy brain can be used to guide the identification of lesion sites that predict language outcome after stroke. However, (a) the level of processing that affects performance on phonologically demanding tasks in patients with damage to either of the identified TMS-guided regions has not yet been investigated in depth; (b) I also need to determine whether TMS-guided SMG or pOp lesions impair phonological more than semantic processing, as predicted by prior TMS studies (e.g., Gough et al., 2005; Hartwigsen et al., 2010a, b; Sliwinska et al., 2015; Hartwigsen et al., 2016); (c) it would be relevant to know whether the functional impairment that arises following damage to either the TMS-guided SMG or pOp regions differs in any measurable way. All these issues will be addressed in my fourth experiment (see next chapter).

5.6 Conclusions

This study showed that the ability to discriminate patients with versus without phonological impairments was substantially better when using the TMS-guided regions than regions derived from: (i) an unguided lesion overlap map of patients with selective

phonological impairments; (ii) a voxel-based statistical comparison of lesion sites in patients with versus without phonological impairments; or (iii) a voxel-based multiple regression analysis that correlated the degree of damage with the severity of phonological impairments. The highest classification accuracy was observed when the TMS-guided SMG and pOp regions were combined with ROI-2 in Sample 1 or ROI-3 in Sample 3.

These findings demonstrate (i) that the seed TMS regions (from previous studies of neurologically-normal individuals) were essential for informing the search for critical lesion sites, (ii) how the integration of results from a variety of lesion-deficit mapping techniques can help to locate all the parts of a distributed cognitive system where brain damage disrupts behaviour, and (iii) the importance of the mid portion of the left superior temporal gyrus (which was part of ROI-2 and ROI-3) for phonological processing.

CHAPTER 6 (Experiment IV):

Investigating the Functional Role of the TMS-guided SMG and pOp Regions

6.1 Summary

Experiment IV sought to characterise in greater detail the functional role played by the TMS-guided SMG and pOp regions. By analysing patients' performance across a collection of tasks that placed varying demands on a wide range of sensory, motor and cognitive functions (such as auditory/visual perception, phonology, semantics and speech articulation), I found that: (i) damage to both TMS-guided regions disrupted the ability to perform tasks that involved phonological processing with and without the need to generate overt speech responses; (ii) the association between damage and impaired performance was significantly greater for phonologically than semantically demanding tasks following lesions to either region; and (iii) there were no noticeable differences in the behaviour of patients with damage to the TMS-guided SMG versus pOp regions. Taken together, these findings suggest that the TMS-guided SMG and pOp regions play a critical role in covert phonological processing, which is consistent with both being part of the dorsal language stream that supports the mapping between sensory and motor speech codes for overt and covert articulation.

6.2 Introduction

The goal of my fourth study was to characterise in greater detail the functional role played by the TMS-guided regions (i.e. SMG and pOp), which could be of particular relevance to gaining a deeper understanding of the results reported in previous chapters as well as findings from other lesion, TMS and fMRI studies of phonological processing.

The original studies of the undamaged brain that identified the seed TMS SMG and pOp regions used in Experiments I and II (i.e. Gough et al., 2005 and Sliwinska et al., 2015), employed a phonological decision task that did not require speech production. Specifically, neurologically-normal participants were asked to make a button press response to indicate whether pairs of visually presented words that had different spellings sounded the same (e.g., jeans - genes). It was found that TMS over SMG or pOp relative to no stimulation caused significant slowing in response times during phonological decisions but not during semantic decisions that involved making button press responses to indicate whether pairs of visually presented words were related in meaning (e.g., gift - present). The increased sensitivity to TMS during phonological contrasted with semantic processing is consistent with ample evidence from prior functional imaging studies that have shown higher activation in SMG and/or pOp for phonological than semantic tasks (Price et al., 1997; Poldrack et al., 1999; Booth et al., 2002; Devlin et al., 2003; McDermott et al., 2003; Seghier et al., 2004; Gitelman et al., 2005; Simard et al., 2013). It is also in line with reports that TMS over SMG or pOp was associated with less efficient performance on phonological relative to semantic decisions, irrespective of input modality (Hartwigsen et al., 2010a, b).

The observation that TMS over SMG or pOp affects phonological decision performance independently of input modality (i.e. visual or auditory) (Hartwigsen et al., 2010a, b) suggests that the recorded behavioural effect is not the consequence of disrupted perceptual processing. Likewise, an explanation in terms of perturbed motor speech production can be excluded because the phonological decisions used in prior TMS studies of SMG and pOp processing involved button press rather than overt speech

responses. There is also prior functional imaging (Crottaz-Herbette et al., 2004; Wimmer et al., 2010) and lesion (Geva et al., 2011) evidence that SMG and pOp support covert articulation after controlling for visual-to-phonological recoding. Together, previous TMS, fMRI and lesion studies of neurologically-normal subjects and brain-damaged patients therefore indicate that parts of SMG and pOp are involved in covert phonological processing.

In this context, I hypothesised that the degree of damage to the TMS-guided SMG or pOp regions would be a significant predictor of inter-patient differences in performance on phonologically demanding tasks more than semantically demanding tasks, irrespective of the stimulus modality (visual or auditory) or whether overt speech production was required. Testing these hypotheses in the samples of patients from Experiments I-III is essential because the lesion studies in those chapters used larger regions of interest (TMS-guided) that involve a combination of white and grey matter. Moreover, phonological processing abilities were measured using speech production tasks (i.e. non-word reading and digit span) rather than phonological decisions with button press responses. It is therefore possible that stroke damage to the TMS-guided regions may have a larger impact on overt than covert phonological processing tasks if some of the grey and white matter included in these regions is necessary for speech articulation or if the demands on phonological processing are not completely matched across tasks (i.e. higher for overt than covert phonological tasks). Likewise, it is possible that stroke damage to the TMS-guided regions may impair semantic as well as phonological tasks if some of the grey and white matter included in these regions is necessary for a type of processing that is common to both phonological and semantic tasks.

Another point that remains to be addressed is whether the effects of damage on behaviour differ for the SMG and pOp regions. In this sense, previous studies on the connectivity profiles of SMG and pOp have shown that these two regions are structurally and functionally densely interconnected (Catani et al., 2005; Makris et al., 2005; Xiang

et al., 2010; Margulies and Petrides, 2013; Martino et al., 2013). Moreover, by using a multiple receptor mapping approach, it has been demonstrated that the distribution of transmitter receptors in SMG is similar to that in pOp (Amunts et al., 2010; Caspers et al., 2013), suggesting the existence of a shared functional substrate between these two regions. Consequently, I predicted that the performance pattern of patients with damage to TMS-guided SMG would closely resemble that of patients with damage to TMS-guided pOp across a wide range of language tasks. Alternatively, the presence of lesion-specific task effects would point to a greater degree of functional differentiation between SMG and pOp than is possible to infer from prior research findings.

6.2.1 Research question

In brief, the current experiment attempted to answer the following research question:

- What is the functional role of the TMS-guided SMG and pOp regions?

6.2.2 Rationale

If the functional impairment caused by damage to the TMS-guided regions is at the level of covert phonological processing, then damage to the TMS-guided regions should predict the ability to perform tasks that involve phonological processing but do not necessitate overt speech production. In the test battery administered to all patients (CAT; Swinburn et al., 2004), the writing-to-dictation task (hear a word and write it down) involves covert phonology (holding a phonological representation of the heard word in memory while the associated orthography is retrieved) because the response is in writing (i.e. it does not require overt speech production). I therefore investigated how well damage to the TMS-guided regions predicted performance on writing-to-dictation after controlling for other types of processing involved in this task including: (i) visual/orthographic processing, (ii) lexical-semantic processing and (iii) hand writing (copying text).

Conversely, if the functional impairment caused by damage to the TMS-guided regions is at the level of overt phonological processing, then damage to the TMS-guided

SMG and pOp regions should predict performance on tasks involving overt speech production (e.g., non-word reading and digit span) but not on tasks that do not require overt speech production (i.e. writing-to-dictation).

In addition, I investigated the degree to which damage to one or the other of the TMS guided regions was associated with performance on semantic tasks that either involved matching heard/seen words to pictures or semantic associations between two pictures. This allowed me to compare the strength of the lesion-deficit mapping for a range of tasks that varied in their demands on phonological and semantic processing.

Finally, I compared the performance of patients with damage to TMS-guided SMG versus pOp across the 27 tasks that comprise the Comprehensive Aphasia Test (CAT).

6.3 Materials and Methods

6.3.1 Participants

To minimise confounds from recovery of language function, only patients who were between 1 and 5 years post-stroke onset were included in the current study (i.e. Sample 1 from Experiments I, II and III).

6.3.2 Task analysis

The strength of the lesion-deficit association was considered for two phonological tasks involving overt speech production (non-word reading and digit span), one phonological task that required covert phonological processing without overt speech production (writing-to-dictation) and five control tasks: (i) visual word-to-picture matching; (ii) auditory word-to-picture matching; (iii) semantic associations of pictures of objects; (iv) reading function words; and (v) copying text. Details of each task can be found in Chapter 2. Figure 6.1 shows (a) an analysis of the levels of processing that might be engaged by each task and (b) how the combination of multiple tasks allows inferences to be made about the functional role played by the TMS-guided regions.

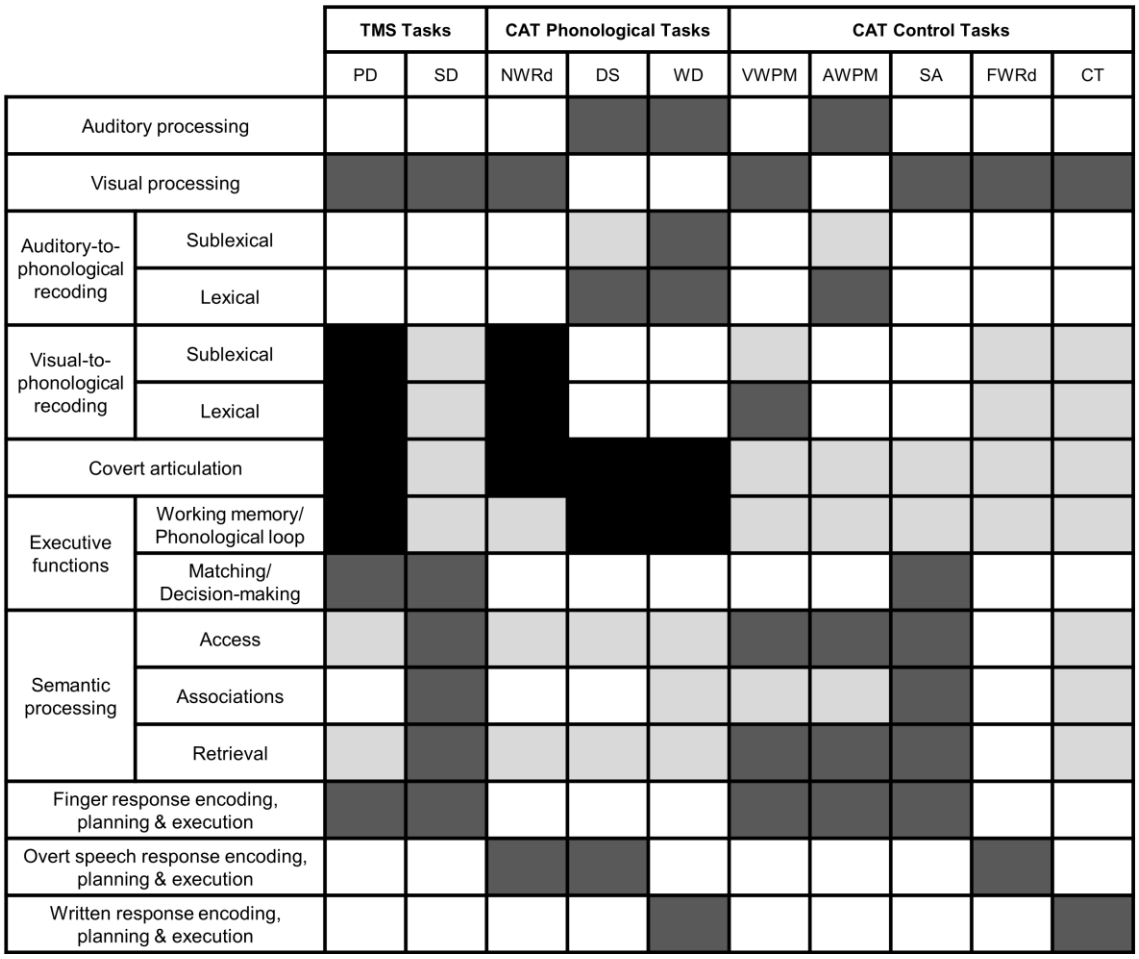

Figure 6.1: Task analysis. The levels of processing hypothesised to be required for completing the TMS phonological (PD) and semantic (SD) decision tasks, and the following tasks from the Comprehensive Aphasia Test (CAT): non-word reading (NWRd), digit span (DS), writing-to-dictation (WD), visual word-to-picture matching (VWPM), auditory word-to-picture matching (AWPM), semantic associations (SA), function word reading (FWRd), and copying text (CT). Black is used to highlight the phonological processes of interest that are shared by the TMS phonological task and at least one of the CAT phonological tasks. Dark grey indicates necessary/explicit processes. Light grey signifies supporting/implicit processes.

6.3.3 Comparing the effect of damage to the TMS-guided regions on covert versus overt phonological processing

Three different multiple regression models were used to test whether the TMS-guided SMG and pOp regions are involved in covert versus overt phonological processing:

(1) The first model included non-word reading scores as the outcome variable and percentage of damage to TMS-guided SMG or pOp (in separate analyses) as the predictor of interest. In addition, the following behavioural covariates of no interest

were included to control for task components unrelated to the function being investigated (i.e. phonological processing): (i) function word reading scores (which are sensitive to the visual/orthographic processing of written words and speech articulation), (ii) visual word-to-picture matching scores (which are sensitive to the visual/orthographic processing of written words and lexical-semantic processing) and (iii) semantic associations scores (which are sensitive to semantic processing).

(2) The second model included digit span scores as the outcome variable and percentage of damage to TMS-guided SMG or pOp (in separate analyses) as the predictor of interest. The behavioural covariates of no interest were: (i) function word reading scores, (ii) auditory word-to-picture matching scores (which are sensitive to the auditory/phonetic processing of spoken words and lexical-semantic processing) and (iii) semantic associations scores.

(3) The third model included writing-to-dictation scores as the outcome variable and percentage of damage to TMS-guided SMG or pOp (in separate analyses) as the predictor of interest. This assessed whether the SMG and pOp regions were required for a task involving covert but not overt phonological processing. The behavioural covariates of no interest were: (i) copying text scores (which are sensitive to hand writing), (ii) auditory word-to-picture matching scores and (iii) semantic associations scores. Of note, performance on the writing-to-dictation task could be aided by accessing the semantic representations of the heard words; particularly, for highly frequent and imageable lexical items.

6.3.4 Comparing the effect of damage to the TMS-guided regions on phonological versus semantic processing abilities

To determine whether the TMS-guided SMG and pOp regions are preferentially involved in phonological rather than semantic processing, the strength of the relationship between the degree of TMS-guided SMG or pOp damage and scores on each of the phonological tasks (i.e. non-word reading, digit span and writing-to-dictation) was compared with that between the degree of TMS-guided SMG or pOp damage and scores

on each of the semantic tasks (i.e. visual/auditory word-to-picture matching and semantic associations).

6.3.5 Testing for differences in the effect of damage to TMS-guided SMG versus pOp

To assess region-specific effects of damage on behaviour, the performance of patients with above-threshold damage to TMS-guided SMG was compared to that of patients with above-threshold damage to TMS-guided pOp across the 27 tasks comprising the Comprehensive Aphasia Test (CAT; for details, see Chapter 2), which systematically vary the demands on a wide range of sensory, motor and cognitive functions (Swinburn et al., 2004). For critical damage thresholds, see Table 4.4 in Chapter 4.

All the statistical analyses described above were conducted in IBM SPSS Statistics for Windows, Version 22.0 (IBM Corp., Armonk, New York, USA).

6.4 Results

6.4.1 Comparing the effect of damage to the TMS guided-regions on covert versus overt phonological processing

Inter-patient differences in covert phonological processing abilities were significantly predicted by the degree of TMS-guided SMG and/or pOp damage ($p < 0.01$) irrespective of output modality (spoken for non-word reading/digit span and written for writing-to-dictation) and even after accounting for the effect of auditory/visual processing, semantic processing, speech articulation and hand writing skills (see Table 6.1).

6.4.2 Comparing the effect of damage to the TMS-guided regions on phonological versus semantic processing abilities

Consistent with the findings from prior TMS studies of neurologically-normal individuals, greater damage to TMS-guided SMG or pOp was significantly associated with poorer performance, and significantly more so for phonologically than semantically demanding tasks ($p < 0.05$ for all comparisons). The only comparison that did not quite

reach statistical significance (p = 0.066) was the difference in the strength of the relationship between digit span versus visual word-to-picture matching performance and damage to the TMS-guided pOp region (see Table 6.2). Plausibly, this occurred because the visual word-to-picture matching task involves orthographic-to-phonological recoding.

Table 6.1: Detrimental effect of damage to the TMS-guided regions on covert and overt phonological processing.

Model	Outcome	Predictors				Covariates
		SMG only	pOp only	SMG & pOp		
		β	*β*	*β*	*β*	
1	NWRd	-0.07***	-0.07***	-0.06**	-0.06**	FWRd + VWPM + SA
2	DS	-0.11***	-0.06***	-0.10***	-0.04**	FWRd + AWPM + SA
3	WD	-0.12***	-0.10***	-0.09***	-0.08***	CT + AWPM + SA

β = unstandardized beta coefficients from the multiple regression analyses indexing the detrimental effect (i.e. minus sign) of TMS-guided SMG and/or pOp damage on phonological processing abilities; ** = statistically significant at $p < 0.01$; *** = statistically significant at $p < 0.001$. See Figure 6.1 legend for key to abbreviations.

6.4.3 Testing for differences in the effect of damage to TMS-guided SMG versus pOp

Performance on all 27 tasks from the CAT was compared in patients with above-threshold damage to the TMS-guided SMG region but not the pOp region (n = 10) versus TMS-guided pOp region but not the SMG region (n = 25). There were no statistically significant differences in the incidence or severity of impairments after Bonferroni correction for multiple comparisons (i.e. $p < 0.002$); see Figure 6.2. Without Bonferroni correction, non-word repetition scores were lower in patients with above-threshold damage to TMS-guided SMG (M = 44.2, SD = 6.4) than pOp (M = 52.3, SD = 8.0; $t(33) = 2.85$, p = 0.008). No other between-group differences in task performance yielded p values smaller than 0.05.

Table 6.2: Preferential involvement of the TMS-guided regions in phonological versus semantic processing.

			SMG		
			AWPM	**VWPM**	**SA**
SMG	**NWRd**	***r***	-0.51*** vs. -0.28***	-0.51*** vs. -0.21*	-0.51*** vs. -0.07
		z	2.68	3.68	4.69
		p	0.007	< 0.001	< 0.001
	DS	***r***	-0.63*** vs. -0.28***	-0.63*** vs. -0.21*	-0.63*** vs. -0.07
		z	4.64	5.29	5.83
		p	< 0.001	< 0.001	< 0.001
	WD	***r***	-0.60*** vs. -0.28***	-0.60*** vs. -0.21*	-0.60*** vs. -0.07
		z	4.29	5.42	5.97
		p	< 0.001	< 0.001	< 0.001
			pOp		
			AWPM	**VWPM**	**SA**
pOp	**NWRd**	***r***	-0.53*** vs. -0.29***	-0.53*** vs. -0.36***	-0.53*** vs. -0.09
		z	2.88	2.27	4.70
		p	0.004	0.024	< 0.001
	DS	***r***	-0.51*** vs. -0.29***	-0.51*** vs. -0.36***	-0.51*** vs. -0.09
		z	2.72	1.84	4.05
		p	0.006	0.066	< 0.001
	WD	***r***	-0.57*** vs. -0.29***	-0.57*** vs. -0.36***	-0.57*** vs. -0.09
		z	3.78	3.11	5.34
		p	< 0.001	0.002	< 0.001

r = correlation coefficient (i.e. Pearson's *r*); *z* = z-score for the difference in correlation strength; * = statistically significant at $p < 0.05$; *** = statistically significant at $p < 0.001$. Nine patients who had missing scores on the writing-to-dictation (WD) and/or semantic associations (SA) tasks were excluded from all correlation analyses (i.e. $N = 145$). See Figure 6.1 legend for key to abbreviations.

6.5 Discussion

Experiment IV investigated the functional role played by the TMS-guided regions. The results showed that damage to TMS-guided SMG or pOp disrupted the ability to perform tasks that involved covert phonological processing with and without the need to produce overt speech responses, such as when reading non-words, repeating digit strings and writing to dictation words and non-words. Moreover, the degree of TMS-guided SMG or pOp damage accounted for a significantly greater proportion of the variability in performance on phonologically demanding tasks relative to semantically demanding tasks. Finally, no evidence was found that the effect of damage to the TMS-guided SMG region differed from that to the TMS-guided pOp region despite comparing

the performance of the patients across a set of tasks that systematically varied the demands on a wide range of sensory, motor and cognitive functions. These findings suggest that the SMG and pOp regions of interest form part of the same phonological processing system, which breaks down following damage to either part.

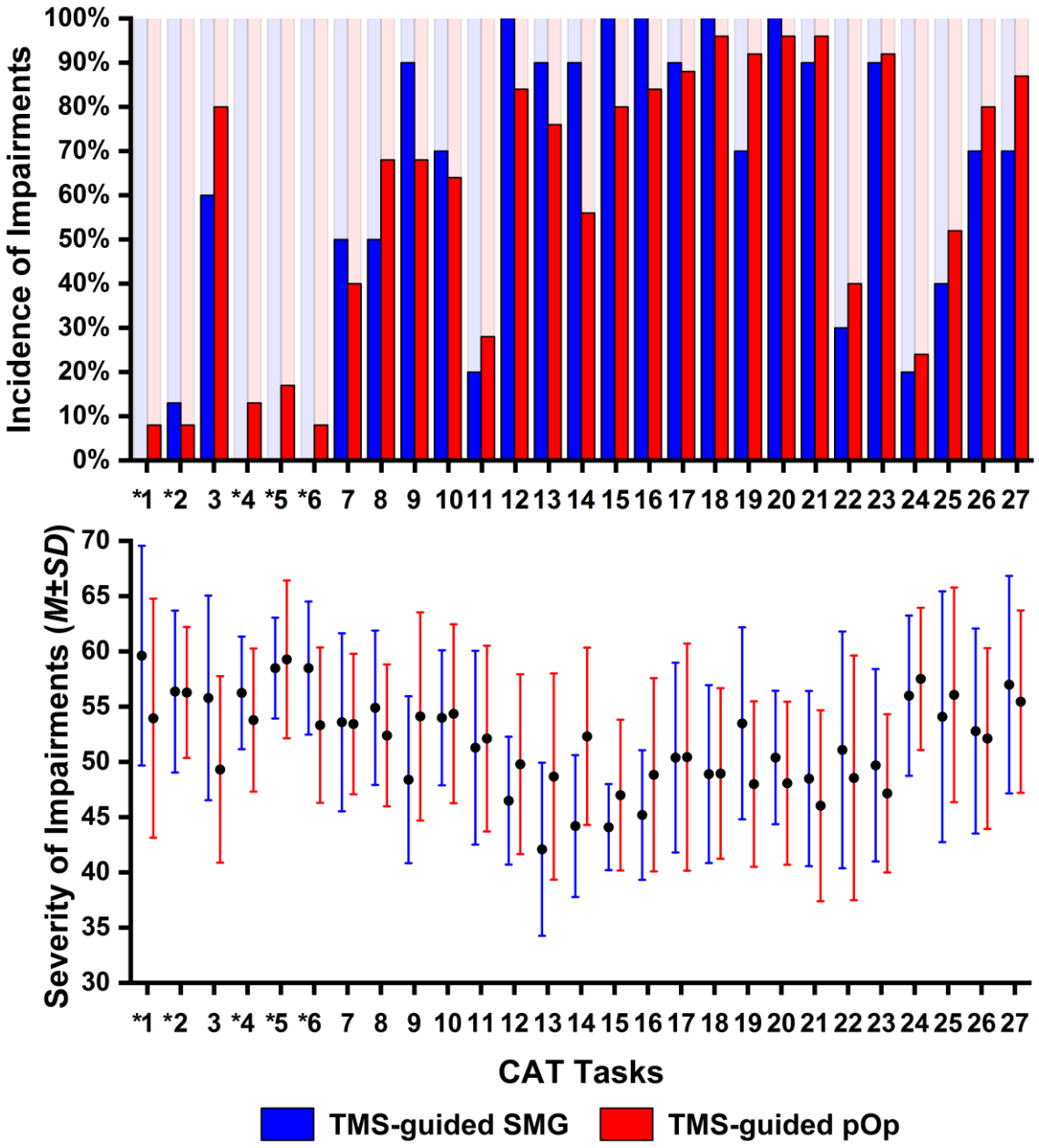

Figure 6.2: Incidence and severity of impairments. Top row: the incidence of impaired performance in patients with above-threshold damage to either TMS-guided SMG (*n* = 10) or TMS-guided pOp (*n* = 25) across all 27 tasks from the CAT is shown. Lighter colours signal spared performance. Bottom row: black circles indicate the group mean T-score for each task, with error bars representing one standard deviation above or below the group mean. * = two patients with above-threshold damage to the TMS-guided SMG region did not complete these tasks from the CAT. See Chapter 2 for which tasks correspond to each number.

6.5.1 Covert versus overt phonological processing

An important advance from the lesion results, over and above previous TMS findings, is that multiple functions were rapidly and safely assessed for the same lesion site. Thus, although the phonological tasks used in the lesion study were not the same as those utilised in the TMS studies (i.e. Gough et al., 2005 and Sliwinska et al., 2015), I was able to investigate how lesions to either the TMS-guided SMG region or the TMS-guided pOp region affected performance across a wide range of language/cognitive tests. This in turn helped to pinpoint the underlying functional impairment. Specifically, by comparing performance on tasks that did and did not require patients to generate a covert phonological representation, I found that the degree of damage to both TMS-guided SMG and pOp was important for explaining inter-patient differences in phonological processing abilities even after factoring out confounds from visual, auditory, semantic and overt speech processing. Furthermore, I found that the detrimental effect of damage to the TMS-guided SMG and pOp regions on performance was not limited to phonological tasks requiring speech production. This renders an interpretation in terms of difficulties arising from a breakdown at the level of the motor execution of speech unlikely. An interpretation in terms of these lesion sites affecting covert phonological processing is, on the other hand, consistent with prior TMS studies of neurologically-normal participants (Gough et al., 2005; Hartwigsen et al., 2010a, b; Sliwinska et al., 2015) which used phonological decision tasks that did not necessitate overt speech responses. Likewise, the fact that damage to the TMS-guided SMG and pOp regions had a negative impact on the ability to write words to dictation is also in agreement with previous functional imaging (e.g., Crottaz-Herbette et al., 2004; Wimmer et al., 2010) and lesion (e.g., Geva et al., 2011) studies associating the left supramarginal gyrus and opercular part of the left inferior frontal gyrus with covert rather than overt phonological processing.

6.5.2 Phonology versus semantics

The observation that greater TMS-guided SMG or pOp damage was more strongly associated with poorer performance on phonologically than semantically demanding tasks mirrors prior findings from TMS (Gough et al., 2005; Hartwigsen et al., 2010a, b; Sliwinska et al., 2015; Hartwigsen et al., 2016) and functional imaging (Price et al., 1997; Poldrack et al., 1999; Booth et al., 2002; Devlin et al., 2003; McDermott et al., 2003; Seghier et al., 2004; Gitelman et al., 2005; Simard et al., 2013) studies of neurologically-normal participants. It is also in keeping with research positing the existence of two functionally dissociable streams for language processing: a dorsal pathway (including SMG and pOp) underlying the mapping from sound to articulation and a ventral pathway supporting the mapping from sound to meaning (e.g., Hickok and Poeppel, 2007; Saur et al., 2008; Kümmerer et al., 2013; Fridriksson et al., 2016).

6.5.3 TMS-guided SMG versus TMS-guided pOp

The comparison of performance over multiple tasks also revealed that the effect of damage to TMS-guided SMG did not differ from the effect of damage to TMS-guided pOp. This is in line with research showing that the receptor fingerprints of SMG and pOp are alike (Amunts et al., 2010; Caspers et al., 2013). It is also consistent with these two areas both being part of the dorsal language pathway that maps sensory and motor phonological representations of speech sounds for covert and overt articulation (Hickok and Poeppel, 2007; Saur et al., 2008; Fridriksson et al., 2016). Since SMG and pOp have been shown to be structurally connected through dorsal white matter tracts (Catani et al., 2005; Makris et al., 2005; Martino et al., 2013; Dick et al., 2014), it is not surprising that stroke lesions that sever these connections impair performance on tasks that load heavily on phonological processing (Rolheiser et al., 2011; Kümmerer et al., 2013). One way that the TMS-guided regions might contribute to covert speech processing is via a dynamic bidirectional information flow that gives rise to a reverberating process (i.e. loop) that allows abstract sensory speech codes to be integrated with their motor counterparts (and vice versa) (Cogan et al., 2014). The assumption of a highly dynamic and interactive

system is in agreement with studies that have examined the spatiotemporal dynamics (Pei et al., 2011; Herman et al., 2013; Liebenthal et al., 2013) and functional connectivity (Xiang et al., 2010; Margulies and Petrides, 2013) of the phonological network.

6.5.4 Limitations

One potential limitation of the current study is that although a fully standardised test battery was utilised to assess the language/cognitive skills of the stroke patients included in Experiment IV, there are no "pure" indices of phonological processing. For instance, both the non-word reading and digit span tasks that defined the presence or absence of phonological processing impairments required overt speech responses. Therefore, impaired performance on either of these tasks could arise at the level of (i) overt articulation, (ii) covert articulation or (iii) a combination of the two. Confounds from speech production could have been avoided if more specific tasks that did not involve any type of speech output such as phonological decisions had been used. Moreover, reaction times measurements (that are not currently available from our test battery) might have indicated that some of the patients that did not show an impairment of accuracy on the non-word reading and digit span tasks were, nonetheless, slower than normal. Here, I capitalised on the opportunities afforded by the Comprehensive Aphasia Test (i.e. a multi-task assessment procedure that has been adopted by many others; see Fyndanis et al., 2017) by performing a detailed a priori task analysis to inform the statistical analyses that allowed me to tease apart effects of interest from uninteresting ones. Therefore, the results of Experiment IV provide a richer picture of the functional role played by the TMS-guided SMG and pOp regions than could possibly be obtained based on the findings from a limited set of very specific behavioural probes.

Another potential weakness is that the expected difference between greater TMS-guided pOp damage and worse performance on digit span than visual word-to-picture matching tasks did not quite reach statistical significance (see Table 6.2). This could have occurred because the demands of the visual word-to-picture matching task on phonological processing were higher than implied by the task analysis (i.e. Figure

6.1). It could also be the case that regions neighbouring pOp were involved in specific aspects of the visual word-to-picture matching task, thereby attenuating the effect of interest (due to co-incidental damage). In addition, disparities in the sensitivity to the behavioural consequences of damage between these two tasks (driven, for instance, by the unequal number of items included) may have concealed the effect of interest. Crucially, however, the reported findings are largely consistent with prior evidence and in line with the predictions laid out at the beginning of the chapter. Future studies will, therefore, need to test if the use of alternative behavioural measures produces similar or different results.

Another potential limitation is that even though a large number of stroke patients (N = 154) were used to examine the function of TMS-guided SMG and TMS-guided pOp, only a relatively small subset of those from the main sample had damage to one or the other of these regions. This could have limited the statistical power to detect some of the effects that were considered (e.g., Button et al., 2013; Lorca-Puls et al., 2018). For example, patients with above-threshold damage to the TMS-guided SMG region tended to obtain lower scores on the non-word repetition task than patients with above-threshold damage to the TMS-guided pOp region. However, the region-specific effect on non-word repetition did not survive when p values were corrected for the number of statistical comparisons conducted (i.e. 27 in total). Beyond potential statistical power issues, however, the results from Experiment IV align well with expectations based on previous research findings (as discussed above).

It, nonetheless, remains possible that TMS-guided SMG damage had a more pronounced impact (albeit not significantly) on the ability to repeat non-words than TMS-guided pOp damage. This might be explained by the fact that the auditory phonological representations of speech sounds have been localised to the mid-posterior portion of the left superior temporal gyrus (STG) (Warren et al., 2005; Desai et al., 2008; McGettigan et al., 2011; Leonard and Chang, 2014; Mesgarani et al., 2014; Evans and Davis, 2015). These auditory representations (in addition to somatosensory ones in SMG) have been

shown to play a critical role in guiding the overt reproduction of incoming streams of aurally presented unfamiliar words (Rogalsky et al., 2015; Leonard et al., 2016; Markiewicz and Bohland, 2016; Basilakos et al., 2017; Behroozmand et al., 2018), which has been incorporated into current neurocomputational models of speech processing (Golfinopoulos et al., 2010; Tourville and Guenther, 2011; Hickok, 2012, 2014). In this context, it is noteworthy that (i) the TMS-guided SMG region included portions of the posterior STG and (ii) due to the spatial proximity of these two areas of the cortex, stroke lesions affecting SMG are more likely to involve mid-posterior STG than lesions to pOp, as indicated by Experiment III.

6.6 Conclusions

This study investigated the type of processing that is affected in patients with damage to the TMS-guided regions. By factoring in performance across a collection of language tasks that systematically varied the demands on phonology, semantics, visual/auditory perception and speech articulation, I found that damage to the TMS-guided regions disrupted phonological processing independently of the involvement of speech production: including reading non-words, repeating digit strings and writing words to dictation. Furthermore, greater damage to the TMS-guided SMG or pOp regions was associated with worse performance on phonologically compared to semantically demanding tasks, suggesting that the TMS-guided regions are preferentially involved in phonological than semantic processing as predicted by previous TMS studies of neurologically-normal individuals. Finally, the behavioural effects of damage were remarkably similar (i.e. no statistically significant differences) for both TMS-guided regions. These findings are, therefore, consistent with prior conclusions that SMG and pOp both play a critical role in covert rather than overt articulation (i.e. covert phonological processing).

CHAPTER 7:

General Discussion

The goal of this thesis was to investigate whether regions derived from previous transcranial magnetic stimulation (TMS) and functional magnetic resonance imaging (fMRI) studies of the undamaged brain could be used to guide the identification of areas of the brain where stroke damage consistently predicts language outcome in individual patients. I additionally compared how well phonological processing abilities were explained by the presence or absence of damage to multiple regions of interest obtained from a range of (new and standard) lesion-deficit mapping approaches. In what follows, I will briefly recapitulate the most relevant findings and their implications for the design of future studies to address unanswered questions and ultimately identify new ways of predicting outcome after stroke.

7.1 The predictive power of the TMS regions is better than that of the fMRI regions

The incidence of phonological processing impairments was predicted significantly better by damage to spherical regions of interest centred on brain sites identified by previous TMS than fMRI studies of phonological processing in neurologically-normal individuals. This is not surprising given that TMS-induced magnetic pulses are used to directly interfere with neural activity in the targeted brain region while behavioural effects are recorded, thereby generating causal evidence for the involvement of the stimulation site in the cognitive function being investigated. In contrast, fMRI measures changes in hemodynamic responses across the whole of the brain as a proxy for the underlying neural activity fluctuations that result from experimental manipulations. The advantage of fMRI relative to TMS is that it can test every cortical area of the brain in the same experiment. The advantage of TMS is that it can test whether specific regions of interest (e.g., those identified by fMRI) are essential for accurate or efficient processing. TMS is therefore an important complement to fMRI because regional activation does not necessarily indicate that the region is essential for task performance. For example, activation could represent (i) false positives (e.g., Eklund et al., 2016; Chen et al., 2018; Takata et al., 2018), (ii) vascular responses that are

downstream to the neural activity (e.g., Sheth et al., 2005), (iii) neural activity from processing that is unrelated to the function of interest (e.g., Christoff, 2012), or (iv) neural activity that is related to a function of interest but can also be supported by other brain regions (Price et al., 1999).

In addition, TMS is useful for localising which part of an activated brain region is most important for the function of interest. This is known as TMS-based functional localisation and was used in the TMS studies that defined the initial regions of interest (i.e. Gough et al., 2005 and Sliwinska et al., 2015). My conclusion that the lesion-deficit mapping was more accurate with TMS than fMRI regions of interest is therefore specifically related to TMS regions identified with functional localisation. Without using functional localisation, the TMS regions would have been exactly the same as the fMRI regions (e.g., Hartwigsen et al., 2010a, b).

Although the predictive power of the TMS sites was better than that of the fMRI sites, the effect of damage to these regions was not entirely consistent across patients. Inter-subject variability in the importance of a brain region for a particular function could arise for multiple reasons. For example, the patients may have had different functional anatomy prior to their stroke or they may have differed in the degree and type of re-learning strategies engaged after their stroke. There are multiple ways to characterise inter-subject variability. I plan to investigate it using fMRI of neurologically-normal participants and examining (i) functional overlap maps that quantify the degree to which activation in specific regions of the brain is consistent or inconsistent across participants (Seghier and Price, 2016) and (ii) data-driven (Kherif et al., 2009; Seghier and Price, 2009) and hypothesis-driven clustering to ask whether more than one processing pathway exists for the same task (Seghier and Price, 2010; Seghier et al., 2014).

If I find that one subset of subjects tends to activate one set of regions (e.g., SMG and pOp) whereas another set of subjects tends to activate another set of regions (e.g., STG and PMC), this would suggest an alternative phonological pathway that could support recovery of language function following loss of pOp or SMG. Such a conclusion

could be further pursued by conducting fMRI studies of patients performing phonological tasks to test (a) whether patients who recovered their phonological processing abilities after damage to one set of phonological areas (e.g., either pOp & SMG or STG & PMC) show increased activation (compared to neurologically-normal controls) in the alternative set of areas and (b) whether the degree of activation in the preserved set of areas is related to the patients' phonological performance in any measurable way.

7.2 The lesion sites underlying persistent phonological impairments involve a combination of grey and white matter

When the borders of the spherical TMS regions of interest were adapted to account for the real extent of the lesions (involving both grey and white matter) in the patients with phonological processing impairments, the prediction accuracy experienced a marked increase even when tested in a completely independent sample of patients. This indicates that persistent phonological impairments may, at least in part, have been caused by damage to the white matter underlying the seed TMS regions. However, in order to ultimately demonstrate that persistent phonological impairments are the consequence of a combination of grey and white matter damage, future lesion studies need to compare the performance of stroke patients with focal damage to the (i) grey matter only, (ii) white matter only or (iii) both. The results might also offer new insights into how TMS affects phonological processing in neurologically-normal individuals. I, nonetheless, anticipate that the likelihood of identifying such patients will be rather small due to the typical extent and shape of vascular lesions.

7.2.1 Inconsistency in the effect of damage to the TMS-guided pOp region

Although the predictive power of the TMS-guided regions was better than that of the original TMS regions, the effect of damage to the TMS-guided regions was still not entirely consistent. Specifically, 10% of patients with damage to the TMS-guided SMG region did not have phonological impairments and 20% of patients with damage to the TMS-guided pOp region did not have phonological impairments. For patients with damage to the TMS-guided SMG region, the consistency of the effect on phonological

processing abilities did not appear to be affected by time post-stroke (1-5 years versus > 5 years). In contrast, damage to the TMS-guided pOp region consistently resulted in phonological impairments in most of the affected patients during the first 5 years post-stroke (i.e. 80%) but inconsistently after 5 years post-stroke (i.e. 50%). I also noted that 8/13 (62%) of patients from Sample 3 (i.e. > 5 years post-stroke) who did not have persistent phonological impairments in the context of TMS-guided pOp damage were more than 9.6 years post-stroke, whereas 12/13 (92%) of those who did have persistent phonological impairments following TMS-guided pOp damage were less than 9.6 years post-stroke. This led me to hypothesise that patients with pOp damage might slowly recover over a period of many years. Indeed, my preliminary findings from an ongoing longitudinal study of patients with TMS-guided pOp (but not TMS-guided SMG) damage suggested that language recovery might start to play a more important role as a function of time since stroke onset (see Experiment II). What is less clear, however, is why language recovery is more apparent in patients with TMS-guided pOp than TMS-guided SMG damage. A tentative explanation might have to do with the strategic location of the TMS-guided SMG region which spatially overlaps with that of a parietotemporal white matter bottleneck that, in stark contrast to a frontal white matter bottleneck, has been shown to contribute to chronic language impairments after left hemispheric stroke (Griffis et al., 2017b).

To identify the lesion and non-lesion factors that may determine who will and will not recover their language abilities in the presence of TMS-guided pOp (or TMS-guided SMG) damage, I plan to conduct a follow-up study involving longitudinal language and fMRI testing of a group of patients with damage to the TMS-guided pOp region and phonological processing impairments. I also plan to search the PLORAS database for patients who perform unexpectedly well on phonologically demanding tasks from the CAT despite damage to the TMS-guided pOp region and who have completed the fMRI paradigm. This would allow me to test whether they show any peri-lesional activation that explains their good phonological performance or whether they engage other brain

regions (see previous section). Mapping the brain regions activated by patients with damage to TMS-guided pOp (or SMG) during phonological processing is essential for explaining inconsistent lesion effects. By combining fMRI data from neurologically-normal controls and stroke patients, I would like to understand how functionally related but anatomically distant regions of the brain communicate with each other during phonological processing and how the phonological network is affected by stroke damage (e.g., Seghier et al., 2010, 2012).

In addition, given that the TMS-guided SMG region has proven to be a very reliable predictor of the incidence of phonological impairments irrespective of time post-stroke, I would like to ascertain its potential for clinical translation by conducting a validation study with another independent sample of patients.

7.3 Phonological processing abilities after stroke are best explained by considering the presence or absence of damage to regions identified by TMS-guided and voxel-based lesion-deficit mapping analyses

The prediction accuracy of the incidence of phonological processing impairments in stroke patients was substantially better when using the TMS-guided regions than an alternative set of regions derived from: (i) an unguided lesion overlap map of patients with selective phonological impairments (ROI-1); (ii) a voxel-based comparison of lesions in patients with versus without phonological impairments (ROI-2); and (iii) a voxel-based multiple regression analysis across patients with and without phonological impairments (ROI-3). Crucially, however, the highest classification accuracy was observed when the TMS-guided regions were combined with ROI-2 or ROI-3. This highlights how the integration of results from different lesion-deficit analyses may help locate the parts of a distributed neural system where brain damage disrupts behaviour. Indeed, TMS, fMRI and univariate voxel-based lesion-deficit mapping studies may reveal the “tip of the iceberg” (i.e. where the most significant group-level effects are observed), but systematic lesion analyses are still required to determine the size and shape of the critical regions

that maximise prediction accuracy (Price et al., 2017; Gajardo-Vidal, Lorca-Puls et al., 2018).

The voxel-based lesion-deficit analyses identified (albeit at an uncorrected voxel-level threshold of $p < 0.001$ one-tailed) two additional regions (i.e. ROI-2 and ROI-3) that both included the mid part of the left superior temporal gyrus. This is in line with the findings from prior studies associating a very similar area located in the mid part of the left superior temporal gyrus with auditory phonological representations, possibly encoded at the level of the syllable (e.g., McGettigan et al., 2011). However, further research is needed to validate the mid part of the left superior temporal gyrus as a relevant predictor of the incidence of phonological processing impairments after stroke.

It would also be informative from a methodological point of view to know why the voxel-based regions were similar but not identical despite both analyses being designed to answer essentially the same question. A potential explanation for the observed discrepancies might be that ROI-2 was identified from a group comparison conducted on binary lesion images (indexing the presence or absence of damage in a categorical manner: only 0s and 1s), whereas ROI-3 was identified from a multiple regression analysis conducted on continuous lesion images (indexing the degree of damage on a continuous scale: from 0 to 1). One way to test this hypothesis would be to replicate each analysis with binary and continuous lesion images and compare the results.

Perhaps more importantly, however, by combining the TMS-guided regions with ROI-2 or ROI-3, I accounted for nearly all the patients in my large cohort who had phonological impairments. This is relevant because it implies that preservation of these regions is likely to support recovery of phonological processing abilities after stroke. This could be tested by monitoring phonological processing abilities, over time, in stroke patients with initial phonological impairments who have lesions that spared the TMS-guided regions, ROI-2 and ROI-3.

7.4 The TMS-guided regions are involved in covert phonological processing

By analysing patients' performance across a wide range of tasks, I was able to replicate and extend the findings from previous TMS and fMRI studies of neurologically-normal individuals. First, I found that the degree of damage to the TMS-guided SMG and/or pOp regions significantly predicted performance on phonologically demanding tasks that did and did not necessitate overt speech responses. Second, I showed that the degree of damage to TMS-guided SMG or TMS-guided pOp accounted for a significantly greater proportion of the variability in performance on phonologically than semantically demanding tasks, which is consistent with prior TMS (Gough et al., 2005; Sliwinska et al., 2015) and fMRI studies of the healthy brain (Devlin et al., 2003; McDermott et al., 2003). Third, I did not find any evidence of differential effects of damage on behaviour for the TMS-guided SMG region compared to the TMS-guided pOp region.

Together, these results suggest that parts of pOp and SMG contribute to an integrated network of regions involved in covert phonological processing, which breaks down following damage to either part. Indeed, Hartwigsen et al. (2016) reported that TMS over both SMG and pOp does not have an additive disruptive effect on phonological processing over and above what is seen after targeting either region alone. Critically, these authors revealed in a subsequent study (Hartwigsen et al., 2017) that focal perturbation of SMG during phonological processing suppresses activity in the whole phonological network including pOp, which explains the absence of an additive detrimental effect on behaviour after concurrently disrupting SMG and pOp. Future lesion studies will need to establish if similar findings are observed irrespective of the particular choice of phonological tasks.

7.5 An 8-step procedure for identifying regions that predict outcome after stroke

Taken together, the results of my experiments suggest that regions that are critical to a function of interest can be identified using the following 8 simple steps:

(1) ***fMRI in neurologically-normal individuals*:** search the whole brain for regions selectively activated by a process of interest (e.g., phonological processing).

(2) ***TMS in neurologically-normal individuals:*** guided by Step 1, use TMS-based functional localisation to identify where disruptive stimulation selectively interferes with the process of interest.

(3) ***Lesion analyses in a large sample of stroke patients:*** guided by Step 1 or 2, stratify patients according to the presence or absence of above-threshold damage to each region of interest.

(4) ***Lesion-outcome analyses in patient subgroups:*** for each subgroup resulting from Step 3, find full extent of grey and/or white matter where the presence or absence of above-threshold damage is most consistently associated with the presence or absence of the deficit of interest (e.g., phonological processing impairments). Use these results to refine the regions of interest.

(5) ***Evaluate prediction/classification accuracy:*** categorise patients according the presence or absence of the deficit of interest following above-threshold damage to each set of regions of interest from Steps 1, 2 & 4. Calculate prediction/classification accuracy (i.e. odds ratio, positive predictive value, negative predictive value, sensitivity and specificity).

(6) ***Lesion-outcome-demographic analyses:*** investigate how the prediction/classification accuracy in Step 5 could be improved by considering the influence of demographic/clinical variables (e.g., age, time post-stroke, handedness, etc.).

(7) ***Validation in a large and independent sample of patients:*** using the criteria from Step 6, calculate how well the prediction/classification accuracy (particularly positive and negative predictive values) replicates in new patients.

(8) ***Clinical translation:*** when lesion sites have high predictive values (positive or negative), new patients can be given a prediction with a confidence rating.

7.6 Conclusions

This thesis has shown how regions of interest derived from previous TMS and fMRI studies of neurologically-normal subjects may be used to guide the identification of lesion sites that consistently predict language outcome after stroke. The proposed methodological procedure can easily be extended to help answer scientifically and clinically relevant questions in other behavioural domains (i.e. apart from language). The findings reported here also indicate that our ability to generate accurate outcome predictions may benefit from the integration of results from multiple lesion-deficit mapping approaches.

References

Abdullaev YG, Bechtereva NP, Melnichuk KV. Neuronal activity of human caudate nucleus and prefrontal cortex in cognitive tasks. Behav Brain Res 1998; 97: 159-77.

Abel S, Weiller C, Huber W, Willmes K, Specht K. Therapy-induced brain reorganization patterns in aphasia. Brain 2015; 138: 1097-112.

Abraham A, Milham MP, Di Martino A, Craddock RC, Samaras D, Thirion B, et al. Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. Neuroimage 2017; 147: 736-45.

Acheson DJ, Hamidi M, Binder JR, Postle BR. A common neural substrate for language production and verbal working memory. J Cogn Neurosci 2011; 23: 1358-67.

Acheson DJ, MacDonald MC. Verbal working memory and language production: Common approaches to the serial ordering of verbal information. Psychol Bull 2009a; 135: 50-68.

Acheson DJ, MacDonald MC. Twisting tongues and memories: Explorations of the relationship between language production and verbal working memory. J Mem Lang 2009b; 60: 329-50.

Aguilar OM, Kerry SJ, Ong Y-H, Callaghan MF, Crinion J, Woodhead ZVJ, et al. Lesion-site-dependent responses to therapy after aphasic stroke. J Neurol Neurosurg Psychiatry 2018. 10.1136/jnnp-2017-317446

Alderson-Day B, Fernyhough C. Inner speech: Development, cognitive functions, phenomenology, and neurobiology. Psychol Bull 2015; 141: 931-65.

Alexander MP, Naeser MA, Palumbo C. Broca's area aphasias: aphasia after lesions including the frontal operculum. Neurology 1990; 40: 353-62.

Allan LM, Rowan EN, Firbank MJ, Thomas AJ, Parry SW, Polvikoski TM, et al. Long term incidence of dementia, predictors of mortality and pathological diagnosis in older stroke survivors. Brain 2011; 134: 3716-27.

Allen L, Mehta S, Andrew McClure J, Teasell R. Therapeutic interventions for aphasia initiated more than six months post stroke: A review of the evidence. Top Stroke Rehabil 2012; 19: 523-35.

Altman DG, Bland JM. Diagnostic tests 1: Sensitivity and specificity. BMJ 1994a; 308: 1552.

Altman DG, Bland JM. Diagnostic tests 2: Predictive values. BMJ 1994b; 309: 102.

Amunts K, Lenzen M, Friederici AD, Schleicher A, Morosan P, Palomero-Gallagher N, et al. Broca's region: Novel organizational principles and multiple receptor mapping. PLoS Biol 2010; 8: e1000489.

Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HBM, Zilles K. Broca's region revisited: Cytoarchitecture and intersubject variability. J Comp Neurol 1999; 412: 319-41.

Arbabshirani MR, Plis S, Sui J, Calhoun VD. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. Neuroimage 2017; 145: 137-65.

Ardila A. A proposed reinterpretation and reclassification of aphasic syndromes. Aphasiology 2010; 24: 363-94.

Arsenault JS, Buchsbaum BR. Distributed neural representations of phonological features during speech perception. J Neurosci 2015; 35: 634-42.

Arthurs OJ, Boniface S. How well do we understand the neural origins of the fMRI BOLD signal? Trends Neurosci 2002; 25: 27-31.

Ashburner J, Friston KJ. Voxel-based morphometry—the methods. Neuroimage 2000; 11: 805-21.

Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005; 26: 839-51.

Assaneo MF, Poeppel D. The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. Sci Adv 2018; 4: eaao3842.

Awh E, Jonides J, Smith EE, Schumacher EH, Koeppe RA, Katz S. Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. Psychol Sci 1996; 7: 25-31.

Awh E, Vogel EK, Oh SH. Interactions between attention and working memory. Neuroscience 2006; 139: 201-8.

Azuar C, Reyes P, Slachevsky A, Volle E, Kinkingnehun S, Kouneiher F, et al. Testing the model of caudo-rostral organization of cognitive control in the human with frontal lesions. Neuroimage 2014; 84: 1053-60.

Baayen RH, Piepenbrock R, Gulikers L. The CELEX lexical database. Philadelphia: Linguistic Data Consortium, University of Pennsylvania; 1995.

Baddeley A. The fractionation of working memory. Proc Natl Acad Sci USA 1996; 93:13468-72.

Baddeley A. The episodic buffer: A new component of working memory? Trends Cogn Sci 2000; 4: 417-23.

Baddeley A. Working memory: Theories, models, and controversies. Annu Rev Psychol 2012; 63: 1-29.

Baddeley AD, Hitch G. Working memory. In: Bower GH, editor. Psychology of learning and motivation. London: Academic Press; 1974. p. 47-89.

Baddeley AD, Hitch GJ. The phonological loop as a buffer store: An update. Cortex 2018. 10.1016/j.cortex.2018.05.015

Baddeley A, Papagno C, Vallar G. When long-term learning depends on short-term storage. J Mem Lang 1988; 27: 586-95.

Baković, Eric. Phonology and phonological theory. In: Goldrick M, Ferreira VS, Miozzo M, editors. The oxford handbook of language production. Oxford: Oxford University Press; 2014. 10.1093/oxfordhb/9780199735471.013.002

Baldo JV, Dronkers NF. The role of inferior parietal and inferior frontal cortex in working memory. Neuropsychology 2006; 20: 529-38.

Baldo JV, Katseff S, Dronkers NF. Brain regions underlying repetition and auditory-verbal short-term memory deficits in aphasia: Evidence from voxel-based lesion symptom mapping. Aphasiology 2012; 26: 338-54.

Barbey AK, Colom R, Solomon J, Krueger F, Forbes C, Grafman J. An integrative architecture for general intelligence and executive function revealed by lesion mapping. Brain 2012; 135: 1154-64.

Barker AT, Jalinous R, Freeston IL. Non-invasive magnetic stimulation of human motor cortex. Lancet 1985; 325: 1106-7.

Basilakos A, Fridriksson J, Rorden C, Behroozmand R, Hanayik T, Naselaris T, et al. Activity associated with speech articulation measured through direct cortical recordings. Brain Lang 2017; 169: 1-7.

Basso A, Lecours AR, Moraschini S, Vanier M. Anatomoclinical correlations of the aphasias as defined through computerized tomography: Exceptions. Brain Lang 1985; 26: 201-29.

Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, et al. Voxel-based lesion–symptom mapping. Nat Neurosci 2003; 6: 448-50.

Behroozmand R, Phillip L, Johari K, Bonilha L, Rorden C, Hickok G, et al. Sensorimotor impairment of speech auditory feedback processing in aphasia. Neuroimage 2018; 165: 102-11.

Benson DF. Fluency in aphasia: Correlation with radioactive scan localization. Cortex 1967; 3: 373-94.

Bernal B, Ardila A. The role of the arcuate fasciculus in conduction aphasia. Brain 2009; 132: 2309-16.

Bernhardt J, Hayward KS, Kwakkel G, Ward NS, Wolf SL, Borschmann K, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: The stroke recovery and rehabilitation roundtable taskforce. Neurorehabil Neural Repair 2017; 31: 793-9.

Berthier ML. Poststroke aphasia: Epidemiology, pathophysiology and treatment. Drugs Aging 2005; 22: 163-82.

Berthier ML, Pulvermüller F. Neuroscience insights improve neurorehabilitation of poststroke aphasia. Nat Rev Neurol 2011; 7: 86-97.

Berthier ML, Pulvermüller F, Dávila G, Casares NG, Gutiérrez A. Drug therapy of post-stroke aphasia: A review of current evidence. Neuropsychol Rev 2011; 21: 302-17.

Binder JR. The Wernicke area: Modern evidence and a reinterpretation. Neurology 2015; 85: 2170-5.

Binder JR. Phoneme perception. In: Hickok G, Small SL, editors. Neurobiology of language. San Diego: Academic Press; 2016. p. 447-61.

Binder JR. Current controversies on Wernicke's area and its role in language. Curr Neurol Neurosci Rep 2017; 17: 58.

Binney RJ, Lambon Ralph MA. Using a combination of fMRI and anterior temporal lobe rTMS to measure intrinsic and induced activation changes across the semantic cognition network. Neuropsychologia 2015; 76: 170-81.

Bitar R, Leung G, Perng R, Tadros S, Moody AR, Sarrazin J, et al. MR pulse sequences: What every radiologist wants to know but is afraid to ask. Radiographics 2006; 26: 513-37.

Bland JM, Altman DG. The odds ratio. BMJ 2000; 320: 1468.

Blank IA, Kiran S, Fedorenko E. Can neuroimaging help aphasia researchers? Addressing generalizability, variability, and interpretability. Cogn Neuropsychol 2017; 34: 377-93.

Blom-Smink MRMA, van de Sandt-Koenderman MWME, Kruitwagen CLJJ, El Hachioui H, Visch-Brink EG, Ribbers GM. Prediction of everyday verbal communicative ability of aphasic stroke patients after inpatient rehabilitation. Aphasiology 2017; 31: 1379-91.

Blumstein SE. Psycholinguistic approaches to the study of syndromes and symptoms of aphasia. In: Hickok G, Small SL, editors. Neurobiology of language. San Diego: Academic Press; 2016. p. 923-33.

Boehme AK, Martin-Schild S, Marshall RS, Lazar RM. Effect of aphasia on acute stroke outcomes. Neurology 2016; 87: 2348-54.

Bohland JW, Bullock D, Guenther FH. Neural representations and mechanisms for the performance of simple speech sequences. J Cogn Neurosci 2010; 22: 1504-29.

Bohland JW, Guenther FH. An fMRI investigation of syllable sequence production. Neuroimage 2006; 32: 821-41.

Booth JR, Burman DD, Meyer JR, Gitelman DR, Parrish TB, Mesulam MM. Modality independence of word comprehension. Hum Brain Mapp 2002; 16: 251-61.

Bouton S, Chambon V, Tyrand R, Guggisberg AG, Seeck M, Karkar S, et al. Focal versus distributed temporal cortex activity for speech sound category assignment. Proc Natl Acad Sci USA 2018; 115: E1299-308.

Boyd LA, Hayward KS, Ward NS, Stinear CM, Rosso C, Fisher RJ, et al. Biomarkers of stroke recovery: Consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. Neurorehabil Neural Repair 2017; 31: 864-76.

Brady MC, Ali M, Fyndanis C, Kambanaros M, Grohmann KK, Laska A-C, et al. Time for a step change? Improving the efficiency, relevance, reliability, validity and transparency of aphasia rehabilitation research through core outcome measures, a common data set and improved reporting criteria. Aphasiology 2014; 28: 1385-92.

Brady MC, Kelly H, Godwin J, Enderby P, Campbell P. Speech and language therapy for aphasia following stroke. Cochrane Database Syst Rev 2016; CD000425.

Brasil-Neto JP, Cohen LG, Panizza M, Nilsson J, Roth BJ, Hallett M. Optimal focal transcranial magnetic activation of the human motor cortex: Effects of coil orientation, shape of the induced current pulse, and stimulus intensity. J Clin Neurophysiol 1992a; 9: 132-6.

Braga-Neto UM, Dougherty ER. Is cross-validation valid for small-sample microarray classification? Bioinformatics 2004; 20: 374-80.

Brasil-Neto JP, McShane LM, Fuhr P, Hallett M, Cohen LG. Topographic mapping of the human motor cortex with magnetic stimulation: Factors affecting accuracy and reproducibility. Electroencephalogr Clin Neurophysiol 1992b; 85: 9-16.

Breitenstein C, Grewe T, Flöel A, Ziegler W, Springer L, Martus P, et al. Intensive speech and language therapy in patients with chronic aphasia after stroke: A randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. Lancet 2017; 389: 1528-38.

Broca P. Nouvelle observation d'aphémie produite par une lésion de la moitié postérieure des deuxième et troisième circonvolutions frontales. Bulletins de la Société Anatomique de Paris 1861a; 36: 398-407.

Broca P. Perte de la parole, ramouissement chronique et destruction partielle du lobe antérieur gauche du cerveau. Bulletins de la Société d'Anthropologie de Paris 1861b; 2: 235-38.

Browman CP, Goldstein L. Articulatory phonology: An overview. Phonetica 1992; 49: 155-80.

Brownsett SLE, Warren JE, Geranmayeh F, Woodhead Z, Leech R, Wise RJS. Cognitive control and its impact on recovery from aphasic stroke. Brain 2014; 137(1): 242-54.

Bruce C, Edmundson A. Letting the CAT out of the bag: A review of the comprehensive aphasia test. Commentary on Howard, Swinburn, and Porter, "Putting the CAT out: What the comprehensive aphasia test has to offer". Aphasiology 2010; 24: 79-93.

Brumberg JS, Krusienski DJ, Chakrabarti S, Gunduz A, Brunner P, Ritaccio AL, et al. Spatio-temporal progression of cortical activity related to continuous overt and covert speech production in a reading task. PLoS One 2016; 11: e0166872.

Buchsbaum BR, D'Esposito M. The search for the phonological store: From loop to convolution. J Cogn Neurosci 2008; 20: 762-78.

Buchsbaum BR, Hickok G, Humphries C. Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. Cogn Sci 2001; 25: 663-78.

Buchsbaum BR, Olsen RK, Koch P, Berman KF. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. Neuron 2005; 48: 687-97.

Buchwald A. Phonetic processing. In: Goldrick M, Ferreira VS, Miozzo M, editors. The oxford handbook of language production. Oxford: Oxford University Press; 2014. 10.1093/oxfordhb/9780199735471.013.013

Buchwald A, Miozzo M. Finding levels of abstraction in speech production: Evidence from sound-production impairment. Psychol Sci 2011; 22: 1113-9.

Burton MW, LoCasto PC, Krebs-Noble D, Gullapalli RP. A systematic investigation of the functional neuroanatomy of auditory and visual phonological processing. Neuroimage 2005; 26: 647-61.

Butler RA, Lambon Ralph MA, Woollams AM. Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. Brain 2014; 137: 3248-66.

Butterworth B, Warrington E. Two routes to repetition: Evidence from a case of 'deep dysphasia'. Neurocase 1995; 1: 55-66.

Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, et al. Power failure: Why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013; 14: 365-76.

Bzdok D, Yeo BTT. Inference in the age of big data: Future perspectives on neuroscience. Neuroimage 2017; 155: 549-64.

Caplan D. Toward a psycholinguistic approach to acquired neurogenic language disorders. Am J Speech Lang Pathol 1993; 2: 59-83.

Caramazza A. The logic of neuropsychological research and the problem of patient classification in aphasia. Brain Lang 1984; 21: 9-20.

Caramazza A, Miceli G, Villa G. The role of the (output) phonological buffer in reading, writing, and repetition. Cogn Neuropsychol 1986; 3: 37-76.

Carlson T, Goddard E, Kaplan DM, Klein C, Ritchie JB. Ghosts in machine learning for cognitive neuroscience: Moving from data to theory. Neuroimage 2018; 180: 88-100.

Carter KN, Anderson CS, Hackett ML, Barber PA, Bonita R. Improved survival after stroke: Is admission to hospital the major explanation? Trend analyses of the Auckland Regional Community Stroke Studies. Cerebrovasc Dis 2007; 23: 162-8.

Caspers S, Geyer S, Schleicher A, Mohlberg H, Amunts K, Zilles K. The human inferior parietal cortex: Cytoarchitectonic parcellation and interindividual variability. Neuroimage 2006; 33: 430-48.

Caspers S, Schleicher A, Bacha-Trams M, Palomero-Gallagher N, Amunts K, Zilles K. Organization of the human inferior parietal lobule based on receptor architectonics. Cereb Cortex 2013; 23: 615-28.

Catani M, Jones DK, ffytche DH. Perisylvian language networks of the human brain. Ann Neurol 2005; 57: 8-16.

Catani M, Mesulam M. The arcuate fasciculus and the disconnection theme in language and aphasia: History and current state. Cortex 2008; 44: 953-61.

Cauquil-Michon C, Flamand-Roze C, Denier C. Borderzone strokes and transcortical aphasia. Curr Neurol Neurosci Rep 2011; 11: 570-7.

Chang EF, Raygor KP, Berger MS. Contemporary model of language organization: An overview for neurosurgeons. J Neurosurg 2015; 122: 250-61.

Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. Categorical speech representation in human superior temporal gyrus. Nat Neurosci 2010; 13: 1428-32.

Charidimou A, Kasselimis D, Varkanitsa M, Selai C, Potagas C, Evdokimidis I. Why is it difficult to predict language impairment and outcome in patients with aphasia after stroke? J Clin Neurol 2014; 10: 75-83.
Chen G, Cox RW, Glen DR, Rajendra JK, Reynolds RC, Taylor PA. A tail of two sides: Artificially doubled false positive rates in neuroimaging due to the sidedness choice with t-tests. Hum Brain Mapp 2018. 10.1002/hbm.24399
Cheng B, Golsari A, Fiehler J, Rosenkranz M, Gerloff C, Thomalla G. Dynamics of regional distribution of ischemic lesions in middle cerebral artery trunk occlusion relates to collateral circulation. J Cereb Blood Flow Metab 2011; 31: 36-40.
Chevillet MA, Jiang X, Rauschecker JP, Riesenhuber M. Automatic phoneme category selectivity in the dorsal auditory stream. J Neurosci 2013; 33: 5208-15.
Christoff K. Undirected thought: Neural determinants and correlates. Brain Res 2012; 1428: 51-9.
Cogan GB, Iyer A, Melloni L, Thesen T, Friedman D, Doyle W, et al. Manipulating stored phonological input during verbal working memory. Nat Neurosci 2017; 20: 279-86.
Cogan GB, Thesen T, Carlson C, Doyle W, Devinsky O, Pesaran B. Sensory–motor transformations for speech occur bilaterally. Nature 2014; 507: 94-8.
Coltheart M. The MRC psycholinguistic database. Q J Exp Psychol A 1981; 33: 497-505.
Coltheart M. Modeling reading: The dual-route approach. In: Snowling MJ, Hulme C, editors. The science of reading: A handbook. Oxford: Blackwell Publishing; 2005. p. 6-23.
Coltheart M, Rastle K, Perry C, Langdon R, Ziegler J. DRC: A dual route cascaded model of visual word recognition and reading aloud. Psychol Rev 2001; 108: 204-56.
Corley M, Brocklehurst PH, Moat HS. Error biases in inner and overt speech: Evidence from tongue twisters. J Exp Psychol Learn Mem Cogn 2011; 37: 162-75.
Correia JM, Jansma BMB, Bonte M. Decoding articulatory features from fMRI responses in dorsal speech regions. J Neurosci 2015; 35: 15015-25.
Coveney PV, Dougherty ER, Highfield RR. Big data need big theory too. Philos Trans A Math Phys Eng Sci 2016; 374: 20160153.
Cowan N. What are the differences between long-term, short-term, and working memory? In: Sossin WS, Lacaille J-C, Castellucci VF, Belleville S, editors. Essence of memory. Prog Brain Res: Elsevier; 2008. p. 323-38.
Cowan N. Working memory and attention in language use. In: Guendouzi J, Loncke F, Williams MJ, editors. The handbook of psycholinguistic and cognitive processes: Perspectives in communication disorders. Hove: Psychology Press; 2010. p. 75-97.
Crary MA, Wertz RT, Deal JL. Classifying aphasias: Cluster analysis of western aphasia battery and boston diagnostic aphasia examination results. Aphasiology 1992; 6: 29-36.
Crinion J, Ashburner J, Leff A, Brett M, Price C, Friston K. Spatial normalization of lesioned brains: Performance evaluation and impact on fMRI analyses. Neuroimage 2007; 37: 866-75.
Crisp J, Lambon Ralph MA. Unlocking the nature of the phonological-deep dyslexia continuum: The keys to reading aloud are in phonology and semantics. J Cogn Neurosci 2006; 18: 348-62.
Croquelois A, Bogousslavsky J. Stroke aphasia: 1,500 consecutive cases. Cerebrovasc Dis 2011; 31: 392-9.
Crottaz-Herbette S, Anagnoson RT, Menon V. Modality effects in verbal working memory: Differential prefrontal and parietal responses to auditory and visual stimuli. Neuroimage 2004; 21: 340-51.

Cui Z, Gong G. The effect of machine learning regression algorithms and sample size on individualized behavioral prediction with functional connectivity features. Neuroimage 2018; 178: 622-37.

Currie S, Hoggard N, Craven IJ, Hadjivassiliou M, Wilkinson ID. Understanding MRI: Basic MR physics for physicians. Postgrad Med J 2013; 89: 209-23.

D'Esposito M, Postle BR. The cognitive neuroscience of working memory. Annu Rev Psychol 2015; 66: 115-42.

Damasio AR, Geschwind N. The neural basis of language. Annu Rev Neurosci 1984; 7: 127-47.

Danelli L, Marelli M, Berlingeri M, Tettamanti M, Sberna M, Paulesu E, et al. Framing effects reveal discrete lexical-semantic and sublexical procedures in reading: An fMRI study. Front Psychol 2015; 6: 1328.

Dayan E, Censor N, Buch ER, Sandrini M, Cohen LG. Noninvasive brain stimulation: From physiology to network dynamics and back. Nat Neurosci 2013; 16: 838-44.

de Haan B, Karnath H-O. A hitchhiker's guide to lesion-behaviour mapping. Neuropsychologia 2018; 115: 5-16.

Deichmann R, Schwarzbauer C, Turner R. Optimisation of the 3D MDEFT sequence for anatomical brain imaging: Technical implications at 1.5 and 3 T. Neuroimage 2004; 21: 757-67.

DeLeon J, Gottesman RF, Kleinman JT, Newhart M, Davis C, Heidler-Gary J, et al. Neural regions essential for distinct cognitive processes underlying picture naming. Brain 2007; 130: 1408-22.

Dell GS, Martin N, Schwartz MF. A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. J Mem Lang 2007; 56: 490-520.

Dell GS, Schwartz MF, Nozari N, Faseyitan O, Branch Coslett H. Voxel-based lesion-parameter mapping: Identifying the neural correlates of a computational model of word production. Cognition 2013; 128: 380-96.

DeMarco AT, Turkeltaub PE. A multivariate lesion symptom mapping toolbox and examination of lesion-volume biases and correction methods in lesion-symptom mapping. Hum Brain Mapp 2018; 39: 4169-82.

DeMarco AT, Wilson SM, Rising K, Rapcsak SZ, Beeson PM. Neural substrates of sublexical processing for spelling. Brain Lang 2017; 164: 118-28.

Demeurisse G, Demol O, Derouck M, de Beuckelaer R, Coekaerts MJ, Capon A. Quantitative study of the rate of recovery from aphasia due to ischemic stroke. Stroke 1980; 11: 455-8.

Deng Z-D, Lisanby SH, Peterchev AV. Electric field depth–focality tradeoff in transcranial magnetic stimulation: Simulation comparison of 50 coil designs. Brain Stimul 2013; 6: 1-13.

Desai R, Liebenthal E, Waldron E, Binder JR. Left posterior temporal regions are sensitive to auditory categorization. J Cogn Neurosci 2008; 20: 1174-88.

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

Devlin JT, Watkins KE. Stimulating language: Insights from TMS. Brain 2007; 130: 610-22.

Dial H, Martin R. Evaluating the relationship between sublexical and lexical processing in speech perception: Evidence from aphasia. Neuropsychologia 2017; 96: 192-212.

Dick AS, Bernal B, Tremblay P. The language connectome: New pathways, new concepts. Neuroscientist 2013; 20: 453-67.

Dickey L, Kagan A, Lindsay MP, Fang J, Rowland A, Black S. Incidence and profile of inpatient stroke-induced aphasia in Ontario, Canada. Arch Phys Med Rehabil 2010; 91: 196-202.

Ding N, Melloni L, Zhang H, Tian X, Poeppel D. Cortical tracking of hierarchical linguistic structures in connected speech. Nat Neurosci 2016; 19: 158-64.

Dronkers NF. A new brain region for coordinating speech articulation. Nature 1996; 384:159-61.

Dronkers NF, Ivanova MV, Baldo JV. What do language disorders reveal about brain–language relationships? From classic models to network approaches. J Int Neuropsychol Soc 2017; 23: 741-54.

Dronkers NF, Plaisant O, Iba-Zizen MT, Cabanis EA. Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. Brain 2007; 130: 1432-41.

Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. Cognition 2004; 92: 145-77.

Du Y, Buchsbaum BR, Grady CL, Alain C. Noise differentially impacts phoneme representations in the auditory and speech motor systems. Proc Natl Acad Sci USA 2014; 111: 7126-31.

Eickhoff SB, Bzdok D, Laird AR, Kurth F, Fox PT. Activation likelihood estimation meta-analysis revisited. Neuroimage 2012; 59: 2349-61.

Eickhoff SB, Constable RT, Yeo BTT. Topographic organization of the cerebral cortex and brain cartography. Neuroimage 2018a; 170: 332-47.

Eickhoff SB, Yeo BTT, Genon S. Imaging-based parcellations of the human brain. Nat Rev Neurosci 2018b; 19: 672-86.

Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. Proc Natl Acad Sci USA 2016; 113: 7900-5.

Ekstrom A. How and when the fMRI BOLD signal relates to underlying neural activity: The danger in dissociation. Brain Res Rev 2010; 62: 233-44.

El Hachioui H, Lingsma HF, van de Sandt-Koenderman MWME, Dippel DWJ, Koudstaal PJ, Visch-Brink EG. Long-term prognosis of aphasia after stroke. J Neurol Neurosurg Psychiatry 2013a; 84: 310-5.

El Hachioui H, Lingsma HF, van de Sandt-Koenderman ME, Dippel DWJ, Koudstaal PJ, Visch-Brink EG. Recovery of aphasia after stroke: A 1-year follow-up study. J Neurol 2013b; 260: 166-71.

El Hachioui H, Sandt-Koenderman MW, Dippel DW, Koudstaal PJ, Visch-Brink EG. The ScreeLing: Occurrence of linguistic deficits in acute aphasia post-stroke. J Rehabil Med 2012; 44: 429-35.

Ellis C, Simpson AN, Bonilha H, Mauldin PD, Simpson KN. The one-year attributable cost of poststroke aphasia. Stroke 2012; 43: 1429-31.

Ellis C, Urban S. Age and aphasia: A review of presence, type, recovery and clinical outcomes. Top Stroke Rehabil 2016; 23: 430-9.

Ellis AW, Young AW. Human cognitive neuropsychology: A textbook with readings. Hove: Psychology Press; 1996.

Engelter ST, Gostynski M, Papa S, Frei M, Born C, Ajdacic-Gross V, et al. Epidemiology of aphasia attributable to first ischemic stroke: Incidence, severity, fluency, etiology, and thrombolysis. Stroke 2006; 37: 1379-84.

Epstein-Peterson Z, Vasconcellos Faria A, Mori S, Hillis AE, Tsapkini K. Relatively normal repetition performance despite severe disruption of the left arcuate fasciculus. Neurocase 2012; 18: 521-6.

Eriksson J, Vogel Edward K, Lansner A, Bergström F, Nyberg L. Neurocognitive architecture of working memory. Neuron 2015; 88: 33-46.

Evans S, Davis MH. Hierarchical organization of auditory and motor representations in speech perception: Evidence from searchlight similarity analysis. Cereb Cortex 2015; 25: 4772-88.

Fan L, Li H, Zhuo J, Zhang Y, Wang J, Chen L, et al. The human Brainnetome atlas: A new brain atlas based on connectional architecture. Cereb Cortex 2016; 26: 3508-26.

Faroqi-Shah Y, Kling T, Solomon J, Liu S, Park G, Braun A. Lesion analysis of language production deficits in aphasia. Aphasiology 2014; 28: 258-77.

Feigin VL, Abajobir AA, Abate KH, Abd-Allah F, Abdulle AM, Abera SF, et al. Global, regional, and national burden of neurological disorders during 1990–2015: A systematic analysis for the Global Burden of Disease Study 2015. Lancet Neurol 2017; 16: 877-97.

Feigin VL, Norrving B, George MG, Foltz JL, Roth GA, Mensah GA. Prevention of stroke: A strategic global imperative. Nat Rev Neurol 2016a; 12: 501-12.

Feigin VL, Roth GA, Naghavi M, Parmar P, Krishnamurthi R, Chugh S, et al. Global burden of stroke and risk factors in 188 countries, during 1990–2013: A systematic analysis for the Global Burden of Disease Study 2013. Lancet Neurol 2016b; 15: 913-24.

Fernández-Miranda JC, Wang Y, Pathak S, Stefaneau L, Verstynen T, Yeh F-C. Asymmetry, connectivity, and segmentation of the arcuate fascicle in the human brain. Brain Struct Funct 2015; 220: 1665-80.

Fiez JA, Damasio H, Grabowski TJ. Lesion segmentation and manual warping to a reference brain: Intra- and interobserver reliability. Hum Brain Mapp 2000; 9: 192-211.

Flinker A, Korzeniewska A, Shestyuk AY, Franaszczuk PJ, Dronkers NF, Knight RT, et al. Redefining the role of Broca's area in speech. Proc Natl Acad Sci USA 2015; 112: 2871.

Flowers HL, Skoretz SA, Silver FL, Rochon E, Fang J, Flamand-Roze C, et al. Poststroke aphasia frequency, recovery, and outcomes: A systematic review and meta-analysis. Arch Phys Med Rehabil 2016; 97: 2188-2201.e8

Forkel SJ, Thiebaut de Schotten M, Dell'Acqua F, Kalra L, Murphy DGM, Williams SCR, et al. Anatomical predictors of aphasia recovery: A tractography study of bilateral perisylvian language networks. Brain 2014; 137: 2027-39.

Formisano E, Kim D-S, Di Salle F, van de Moortele P-F, Ugurbil K, Goebel R. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 2003; 40: 859-69.

Frank RJ, Damasio H, Grabowski TJ. Brainvox: An interactive, multimodal visualization and analysis system for neuroanatomical imaging. Neuroimage 1997; 5: 13-30.

Franklin S, Harhen D, Hayes M, Demos Mc Manus S, Pollock A. Top 10 research priorities relating to aphasia following stroke. Aphasiology 2018; 32: 1388-95.

Fridriksson J, Bonilha L, Rorden C. Severe Broca's aphasia without Broca's area damage. Behav Neurol 2007; 18: 237-8.

Fridriksson J, den Ouden D-B, Hillis AE, Hickok G, Rorden C, Basilakos A, et al. Anatomy of aphasia revisited. Brain 2018; 141: 848-62.

Fridriksson J, Fillmore P, Guo D, Rorden C. Chronic Broca's aphasia is caused by damage to Broca's and Wernicke's areas. Cereb Cortex 2015; 25: 4689-96.
Fridriksson J, Guo D, Fillmore P, Holland A, Rorden C. Damage to the anterior arcuate fasciculus predicts non-fluent speech production in aphasia. Brain 2013; 136: 3451-60.
Fridriksson J, Kjartansson O, Morgan PS, Hjaltason H, Magnusdottir S, Bonilha L, et al. Impaired speech repetition and left parietal lobe damage. J Neurosci 2010; 30: 11057-61.
Fridriksson J, Yourganov G, Bonilha L, Basilakos A, Den Ouden D-B, Rorden C. Revealing the dual streams of speech processing. Proc Natl Acad Sci USA 2016; 113: 15108-13.
Friederici AD. The brain basis of language processing: From structure to function. Physiol Rev 2011; 91: 1357-92.
Friederici AD, Gierhan SME. The language network. Curr Opin Neurobiol 2013; 23: 250-4.
Fuertinger S, Horwitz B, Simonyan K. The functional connectome of speech control. PLoS Biol 2015; 13: e1002209.
Fukuda M, Poplawsky AJ, Kim SG. Submillimeter-resolution fMRI: Toward understanding local neural processing. In: Masamoto K, Hirase H, Yamada K, editors. New horizons in neurovascular coupling: A bridge between brain circulation and neural plasticity. Prog Brain Res: Elsevier; 2016. p. 123-52.
Fyndanis V, Lind M, Varlokosta S, Kambanaros M, Soroli E, Ceder K, et al. Cross-linguistic adaptations of the comprehensive aphasia test: Challenges and solutions. Clin Linguist Phon 2017; 31: 697-710.
Gajardo-Vidal A, Lorca-Puls DL, Crinion JT, White J, Seghier ML, Leff AP, et al. How distributed processing produces false negatives in voxel-based lesion-deficit analyses. Neuropsychologia 2018; 115: 124-33.
Galantucci S, Tartaglia MC, Wilson SM, Henry ML, Filippi M, Agosta F, et al. White matter damage in primary progressive aphasias: A diffusion tensor tractography study. Brain 2011; 134: 3011-29.
Galluzzi C, Bureca I, Guariglia C, Romani C. Phonological simplifications, apraxia of speech and the interaction between phonological and phonetic processing. Neuropsychologia 2015; 71: 64-83.
Gathercole SE, Baddeley AD. Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. J Mem Lang 1989; 28: 200-13.
Gelfand JR, Bookheimer SY. Dissociating neural mechanisms of temporal sequencing and processing phonemes. Neuron 2003; 38: 831-42.
Genon S, Reid A, Langner R, Amunts K, Eickhoff SB. How to characterize the function of a brain region. Trends Cogn Sci 2018; 22: 350-64.
Geranmayeh F, Chau TW, Wise RJS, Leech R, Hampshire A. Domain-general subregions of the medial prefrontal cortex contribute to recovery of language after stroke. Brain 2017; 140: 1947-58.
Geranmayeh F, Leech R, Wise RJS. Network dysfunction predicts speech production after left hemisphere stroke. Neurology 2016; 86: 1296-305.
Geschwind N. Disconnexion syndromes in animals and man. I. Brain 1965a; 88: 237-94.
Geschwind N. Disconnexion syndromes in animals and man. II. Brain 1965b; 88: 585-644.
Geschwind N. The organization of language and the brain. Science 1970; 170: 940-4.
Geschwind N. Aphasia. N Engl J Med 1971; 284: 654-6.

Geva S, Baron J-C, Jones PS, Price CJ, Warburton EA. A comparison of VLSM and VBM in a cohort of patients with post-stroke aphasia. Neuroimage Clin 2012; 1: 37-47.

Geva S, Jones PS, Crinion JT, Price CJ, Baron J-C, Warburton EA. The neural correlates of inner speech defined by voxel-based lesion–symptom mapping. Brain 2011; 134: 3071-82.

Gitelman DR, Nobre AC, Sonty S, Parrish TB, Mesulam MM. Language network specializations: An analysis with parallel task designs and functional magnetic resonance imaging. Neuroimage 2005; 26: 975-85.

Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PMM. The diagnostic odds ratio: A single indicator of test performance. J Clin Epidemiol 2003; 56: 1129-35.

Gläscher J, Adolphs R, Damasio H, Bechara A, Rudrauf D, Calamia M, et al. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. Proc Natl Acad Sci USA 2012; 109: 14681-86.

Glasser MF, Coalson TS, Robinson EC, Hacker CD, Harwell J, Yacoub E, et al. A multi-modal parcellation of human cerebral cortex. Nature 2016; 536: 171-78.

Goense J, Bohraus Y, Logothetis NK. fMRI at high spatial resolution: Implications for BOLD-models. Front Comput Neurosci 2016; 10: 66.

Goense JBM, Logothetis NK. Neurophysiology of the BOLD fMRI signal in awake monkeys. Curr Biol 2008; 18: 631-40.

Gold BT, Balota DA, Kirchhoff BA, Buckner RL. Common and dissociable activation patterns associated with controlled semantic and phonological processing: Evidence from fMRI adaptation. Cereb Cortex 2005; 15: 1438-50.

Goldrick M. Phonological processing: The retrieval and encoding of word form information in speech production. In: Goldrick M, Ferreira VS, Miozzo M, editors. The oxford handbook of language production. Oxford: Oxford University Press; 2014. 10.1093/oxfordhb/9780199735471.013.015

Goldrick M, Blumstein SE. Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. Lang Cogn Process 2006; 21: 649-83.

Goldrick M, Rapp B. A restricted interaction account (RIA) of spoken word production: The best of both worlds. Aphasiology 2002; 16: 20-55.

Goldrick M, Rapp B. Lexical and post-lexical phonological representations in spoken production. Cognition 2007; 102: 219-60.

Golfinopoulos E, Tourville JA, Bohland JW, Ghosh SS, Nieto-Castanon A, Guenther FH. fMRI investigation of unexpected somatosensory feedback perturbation during speech. Neuroimage 2011; 55: 1324-38.

Golfinopoulos E, Tourville JA, Guenther FH. The integration of large-scale neural network modeling and functional brain imaging in speech motor control. Neuroimage 2010; 52: 862-74.

Goodglass H, Kaplan E, Barresi B, editors. The assessment of aphasia and related disorders. Philadelphia: Lippincott Williams & Wilkins; 2001.

Goranskaya D, Kreitewolf J, Mueller JL, Friederici AD, Hartwigsen G. Fronto-parietal contributions to phonological processes in successful artificial grammar learning. Front Hum Neurosci 2016; 10: 551.

Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE. Generation and evaluation of a cortical area parcellation from resting-state correlations. Cereb Cortex 2016; 26: 288-303.

Gorelick PB, Scuteri A, Black SE, DeCarli C, Greenberg SM, Iadecola C, et al. Vascular contributions to cognitive impairment and dementia. Stroke 2011; 42: 2672-713.

Gorno-Tempini ML, Dronkers NF, Rankin KP, Ogar JM, Phengrasamy L, Rosen HJ, et al. Cognition and anatomy in three variants of primary progressive aphasia. Ann Neurol 2004; 55: 335-46.

Gough PM, Nobre AC, Devlin JT. Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. J Neurosci 2005; 25: 8010-6.

Gow DW. The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. Brain Lang 2012; 121: 273-88.

Graham JR, Pereira S, Teasell R. Aphasia and return to work in younger stroke survivors. Aphasiology 2011; 25: 952-60.

Graves WW, Desai R, Humphries C, Seidenberg MS, Binder JR. Neural systems for reading aloud: A multiparametric approach. Cereb Cortex 2010; 20: 1799-815.

Graves WW, Grabowski TJ, Mehta S, Gupta P. The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. J Cogn Neurosci 2008; 20: 1698-710.

Griffis JC, Nenert R, Allendorfer JB, Vannest J, Holland S, Dietz A, et al. The canonical semantic network supports residual language function in chronic post-stroke aphasia. Hum Brain Mapp 2017a; 38: 1636-58.

Griffis JC, Nenert R, Allendorfer JB, Szaflarski JP. Damage to white matter bottlenecks contributes to language impairments after left hemispheric stroke. Neuroimage Clin 2017b; 14: 552-65.

Guenther FH, Hickok G. Neural models of motor speech control. In: Hickok G, Small SL, editors. Neurobiology of language. San Diego: Academic Press; 2016. p. 725-40.

Guenther FH, Vladusich T. A neural theory of speech acquisition and production. J Neurolinguistics 2012; 25: 408-22.

Gupta P, MacWhinney B. Vocabulary acquisition and verbal short-term memory: Computational and neural bases. Brain Lang 1997; 59: 267-333.

Halai AD, Woollams AM, Lambon Ralph MA. Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. Cortex 2017; 86: 275-89.

Halai AD, Woollams AM, Lambon Ralph MA. Predicting the pattern and severity of chronic post-stroke language deficits from functionally-partitioned structural lesions. Neuroimage Clin 2018; 19: 1-13.

Han Z, Ma Y, Gong G, Huang R, Song L, Bi Y. White matter pathway supporting phonological encoding in speech production: A multi-modal imaging study of brain damage patients. Brain Struct Funct 2016; 221: 577-89.

Handwerker DA, Gonzalez-Castillo J, D'Esposito M, Bandettini PA. The continuing challenge of understanding and modeling hemodynamic variation in fMRI. Neuroimage 2012; 62: 1017-23.

Hanley JR, Kay J. An effect of imageability on the production of phonological errors in auditory repetition. Cogn Neuropsychol 1997; 14: 1065-84.

Hanley JR, Kay J, Edwards M. Imageability effects, phonological errors, and the relationship between auditory repetition and picture naming: Implications for models of auditory repetition. Cogn Neuropsychol 2002; 19: 193-206.

Harrison RV, Harel N, Panesar J, Mount RJ. Blood capillary distribution correlates with hemodynamic-based functional imaging in cerebral cortex. Cereb Cortex 2002; 12: 225-33.

Hartwigsen G. Flexible redistribution in cognitive networks. Trends Cogn Sci 2018; 22: 687-98.

Hartwigsen G, Baumgaertner A, Price CJ, Koehnke M, Ulmer S, Siebner HR. Phonological decisions require both the left and right supramarginal gyri. Proc Natl Acad Sci USA 2010a; 107: 16494-9.

Hartwigsen G, Bzdok D, Klein M, Wawrzyniak M, Stockert A, Wrede K, et al. Rapid short-term reorganization in the language network. eLife 2017; 6: e25964.

Hartwigsen G, Price CJ, Baumgaertner A, Geiss G, Koehnke M, Ulmer S, et al. The right posterior inferior frontal gyrus contributes to phonological word decisions in the healthy brain: Evidence from dual-site TMS. Neuropsychologia 2010b; 48: 3155-63.

Hartwigsen G, Saur D, Price CJ, Ulmer S, Baumgaertner A, Siebner HR. Perturbation of the left inferior frontal gyrus triggers adaptive plasticity in the right homologous area during speech production. Proc Natl Acad Sci USA 2013; 110: 16402-7.

Hartwigsen G, Weigel A, Schuschan P, Siebner HR, Weise D, Classen J, et al. Dissociating parieto-frontal networks for phonological and semantic word decisions: A condition-and-perturb TMS study. Cereb Cortex 2016; 26: 2590-601.

Hastie T, Rosset S, Tibshirani R, Zhu J. The entire regularization path for the support vector machine. J Mach Learn Res 2004; 5: 1391-415.

Haufe S, Meinecke F, Görgen K, Dähne S, Haynes J-D, Blankertz B, et al. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage 2014; 87: 96-110.

Hay SI, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017; 390: 1260-344.

Heeger DJ, Ress D. What does fMRI tell us about neuronal activity? Nat Rev Neurosci 2002; 3: 142-51.

Heilman KM. Aphasia and the diagram makers revisited: An update of information processing models. J Clin Neurol 2006; 2: 149-62.

Heim S, Opitz B, Müller K, Friederici AD. Phonological processing during language production: fMRI evidence for a shared production-comprehension network. Brain Res Cogn Brain Res 2003; 16: 285-96.

Henry ML, Wilson SM, Babiak MC, Mandelli ML, Beeson PM, Miller ZA, et al. Phonological processing in primary progressive aphasia. J Cogn Neurosci 2016; 28: 210-22.

Henseler I, Regenbrecht F, Obrig H. Lesion correlates of patholinguistic profiles in chronic aphasia: Comparisons of syndrome-, modality- and symptom-level assessment. Brain 2014; 137: 918-30.

Herman AB, Houde JF, Vinogradov S, Nagarajan SS. Parsing the phonological loop: Activation timing in the dorsal speech stream determines accuracy in speech reproduction. J Neurosci 2013; 33: 5439-53.

Herwig U, Abler B, Schönfeldt-Lecuona C, Wunderlich A, Grothe J, Spitzer M, et al. Verbal storage in a premotor–parietal network: Evidence from fMRI-guided magnetic stimulation. Neuroimage 2003; 20: 1032-41.

Hickok G. The functional neuroanatomy of language. Phys Life Rev 2009; 6: 121-43.

Hickok G. Computational neuroanatomy of speech production. Nat Rev Neurosci 2012; 13: 135-45.

Hickok G. The architecture of speech production and the role of the phoneme in speech processing. Lang Cogn Neurosci 2014; 29: 2-20.

Hickok G, Poeppel D. The cortical organization of speech processing. Nat Rev Neurosci 2007; 8: 393-402.
Hickok G, Rogalsky C, Chen R, Herskovits EH, Townsley S, Hillis AE. Partially overlapping sensorimotor networks underlie speech praxis and verbal short-term memory: Evidence from apraxia of speech following acute stroke. Front Hum Neurosci 2014; 8: 649.
Hilari K, Northcott S. “Struggling to stay connected”: Comparing the social relationships of healthy older people and people with stroke and aphasia. Aphasiology 2017; 31: 674-87.
Hillis AE. Aphasia: Progress in the last quarter of a century. Neurology 2007; 69: 200-13.
Hillis AE, Beh YY, Sebastian R, Breining B, Tippett DC, Wright A, et al. Predicting recovery in acute poststroke aphasia. Ann Neurol 2018; 83: 612-22.
Hillis AE, Caramazza A. Mechanisms for accessing lexical representations for output: Evidence from a category-specific semantic deficit. Brain Lang 1991; 40: 106-44.
Hillis AE, Work M, Barker PB, Jacobs MA, Breese EL, Maurer K. Re-examining the brain regions crucial for orchestrating speech articulation. Brain 2004; 127: 1479-87.
Hoffmann M, Chen R. The spectrum of aphasia subtypes and etiology in subacute stroke. J Stroke Cerebrovasc Dis 2013; 22: 1385-92.
Holland A, Fromm D, Forbes M, MacWhinney B. Long-term recovery in stroke accompanied by aphasia: A reconsideration. Aphasiology 2017; 31: 152-65.
Honea R, Crow TJ, Passingham D, Mackay CE. Regional deficits in brain volume in schizophrenia: A meta-analysis of voxel-based morphometry studies. Am J Psychiatry 2005; 162: 2233-45.
Hope TMH, Leff AP, Prejawa S, Bruce R, Haigh Z, Lim L, et al. Right hemisphere structural adaptation and changing language skills years after left hemisphere stroke. Brain 2017; 140: 1718-28.
Hope TMH, Leff AP, Price CJ. Predicting language outcomes after stroke: Is structural disconnection a useful predictor? Neuroimage Clin 2018; 19: 22-9.
Hope TMH, Parker Jones Ō, Grogan A, Crinion J, Rae J, Ruffle L, et al. Comparing language outcomes in monolingual and bilingual stroke patients. Brain 2015; 138: 1070-83.
Hope TMH, Prejawa S, Parker Jones Ō, Oberhuber M, Seghier ML, Green DW, et al. Dissecting the functional anatomy of auditory word repetition. Front Hum Neurosci 2014; 8: 246.
Hope TMH, Seghier ML, Leff AP, Price CJ. Predicting outcome and recovery after stroke with lesions extracted from MRI images. Neuroimage Clin 2013; 2: 424-33.
Houde J, Nagarajan S. Speech production as state feedback control. Front Hum Neurosci 2011; 5: 82.
Howard D, Nickels L. Separating input and output phonology: Semantic, phonological, and orthographic effects in short-term memory impairment. Cogn Neuropsychol 2005; 22: 42-77.
Howard D, Swinburn K, Porter G. Putting the CAT out: What the comprehensive aphasia test has to offer. Aphasiology 2010a; 24: 56-74.
Howard D, Swinburn K, Porter G. The CAT is now out: A response to the commentaries. Aphasiology 2010b; 24(1): 94-8.
Howe T, Davidson B, Worrall L, Hersh D, Ferguson A, Sherratt S, et al. ‘You needed to rehab … families as well’: Family members’ own goals for aphasia rehabilitation. Int J Lang Commun Disord 2012; 47: 511-21.

Huettel SA, Song AW, McCarthy G, editors. Functional magnetic resonance imaging. Sunderland: Sinauer Associates; 2014.
Huys QJM, Maia TV, Frank MJ. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci 2016; 19: 404-13.
Idsardi WJ, Monahan PJ. Phonology. In: Hickok G, Small SL, editors. Neurobiology of language. San Diego: Academic Press; 2016. p. 141-51.
Ingles JL, Mate-Kole CC, Connolly JF. Evidence for multiple routes of speech production in a case of fluent aphasia. Cortex 1996; 32: 199-219.
Inoue K, Madhyastha T, Rudrauf D, Mehta S, Grabowski T. What affects detectability of lesion–deficit relationships in lesion studies? Neuroimage Clin 2014; 6: 388-97.
Ionta S, Heydrich L, Lenggenhager B, Mouthon M, Fornari E, Chapuis D, et al. Multisensory mechanisms in temporo-parietal cortex support self-location and first-person perspective. Neuron 2011; 70: 363-74.
Isaksson A, Wallman M, Göransson H, Gustafsson MG. Cross-validation and bootstrapping are unreliable in small sample classification. Pattern Recognit Lett 2008; 29: 1960-5.
Jacquemot C, Dupoux E, Bachoud-Lévi A-C. Breaking the mirror: Asymmetrical disconnection between the phonological input and output codes. Cogn Neuropsychol 2007; 24: 3-22.
Jacquemot C, Scott SK. What is the relationship between phonological short-term memory and speech processing? Trends Cogn Sci 2006; 10: 480-6.
Janssen RJ, Mourão-Miranda J, Schnack HG. Making individual prognoses in psychiatry using neuroimaging and machine learning. Biol Psychiatry Cogn Neurosci Neuroimaging 2018; 3: 798-808.
Jefferies E, Sage K, Lambon Ralph MA. Do deep dyslexia, dysphasia and dysgraphia share a common phonological impairment? Neuropsychologia 2007; 45: 1553-70.
Jobard G, Crivello F, Tzourio-Mazoyer N. Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. Neuroimage 2003; 20: 693-712.
Jones DK, Knösche TR, Turner R. White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. Neuroimage 2013; 73: 239-54.
Karnath H-O, Fruhmann Berger M, Küker W, Rorden C. The anatomy of spatial neglect based on voxelwise statistical analysis: A study of 140 patients. Cereb Cortex 2004; 14: 1164-72.
Karnath H-O, Rennig J. Investigating structure and function in the healthy human brain: Validity of acute versus chronic lesion-symptom mapping. Brain Struct Funct 2017; 222: 2059-70.
Karnath H-O, Sperber C, Rorden C. Mapping human brain lesions and their functional consequences. Neuroimage 2018; 165: 180-9.
Kasselimis DS, Simos PG, Peppas C, Evdokimidis I, Potagas C. The unbridged gap between clinical diagnosis and contemporary research on aphasia: A short discussion on the validity and clinical utility of taxonomic categories. Brain Lang 2017; 164: 63-7.
Katzir T, Misra M, Poldrack RA. Imaging phonology without print: Assessing the neural correlates of phonemic awareness using fMRI. Neuroimage 2005; 27: 106-15.
Kazanina N, Bowers JS, Idsardi W. Phonemes: Lexical access and beyond. Psychon Bull Rev 2018; 25: 560-85.
Kertesz A, Lau WK, Polk M. The structural determinants of recovery in Wernicke′s aphasia. Brain Lang 1993; 44: 153-64.
Kherif F, Josse G, Seghier ML, Price CJ. The main sources of intersubject variability in neuronal activation for reading aloud. J Cogn Neurosci 2009; 21: 654-68.

Kim S-G, Jin T, Fukuda M. Spatial resolution of fMRI techniques. In: Ulmer S, Jansen O, editors. fMRI: Basics and clinical applications. Heidelberg: Springer; 2010. p. 15-21.
Kim S-G, Ogawa S. Biophysical and physiological origins of blood oxygenation level-dependent fMRI signals. J Cereb Blood Flow Metab 2012; 32: 1188-206.
Kimberg DY, Coslett HB, Schwartz MF. Power in voxel-based lesion-symptom mapping. J Cogn Neurosci 2007; 19: 1067-80.
Kiran S. What is the nature of poststroke language recovery and reorganization? ISRN Neurol 2012; 2012: 786872.
Kircher T, Nagels A, Kirner-Veselinovic A, Krach S. Neural correlates of rhyming vs. lexical and semantic fluency. Brain Res 2011; 1391: 71-80.
Klöppel S, Abdulkadir A, Jack CR, Koutsouleris N, Mourão-Miranda J, Vemuri P. Diagnostic neuroimaging across diseases. Neuroimage 2012; 61: 457-63.
Knecht S, Dräger B, Deppe M, Bobe L, Lohmann H, Flöel A, et al. Handedness and hemispheric language dominance in healthy humans. Brain 2000; 123: 2512-8.
Knott R, Patterson K, Hodges JR. The role of speech production in auditory-verbal short-term memory: Evidence from progressive fluent aphasia. Neuropsychologia 2000; 38: 125-42.
Koenigs M, Acheson DJ, Barbey AK, Solomon J, Postle BR, Grafman J. Areas of left perisylvian cortex mediate auditory–verbal short-term memory. Neuropsychologia 2011; 49: 3612-9.
Krieger-Redwood K, Gaskell MG, Lindsay S, Jefferies E. The selective role of premotor cortex in speech perception: A contribution to phoneme judgements but not speech comprehension. J Cogn Neurosci 2013; 25: 2179-88.
Kriegeskorte N, Douglas PK. Cognitive computational neuroscience. Nat Neurosci 2018; 21: 1148-60.
Krishnamurthi RV, Moran AE, Feigin VL, Barker-Collo S, Norrving B, Mensah GA, et al. Stroke prevalence, mortality and disability-adjusted life years in adults aged 20-64 years in 1990-2013: Data from the Global Burden of Disease 2013 Study. Neuroepidemiology 2015; 45: 190-202.
Kümmerer D, Hartwigsen G, Kellmeyer P, Glauche V, Mader I, Klöppel S, et al. Damage to ventral and dorsal language pathways in acute aphasia. Brain 2013; 136: 619-29.
Laganaro M. Patterns of impairments in AOS and mechanisms of interaction between phonological and phonetic encoding. J Speech Lang Hear Res 2012; 55: S1535-43.
Lambon Ralph MA, Jefferies E, Patterson K, Rogers TT. The neural and computational bases of semantic cognition. Nat Rev Neurosci 2017; 18: 42-55.
Lametti DR, Nasir SM, Ostry DJ. Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. J Neurosci 2012; 32: 9351-8.
Lancaster JL, Tordesillas-Gutiérrez D, Martinez M, Salinas F, Evans A, Zilles K, et al. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum Brain Mapp 2007; 28: 1194-205.
Laska AC, Hellblom A, Murray V, Kahan T, Von Arbin M. Aphasia in acute stroke and relation to outcome. J Intern Med 2001; 249: 413-22.
Lazar RM, Boehme AK. Aphasia as a predictor of stroke outcome. Curr Neurol Neurosci Rep 2017; 17: 83.
Lazar RM, Mohr JP. Revisiting the contributions of Paul Broca to the study of aphasia. Neuropsychol Rev 2011; 21: 236-9.

Lee Y-S, Turkeltaub P, Granger R, Raizada RDS. Categorical speech processing in Broca's area: An fMRI study using multivariate pattern-based analysis. J Neurosci 2012; 32: 3942-8.

Leff AP, Schofield TM, Crinion JT, Seghier ML, Grogan A, Green DW, et al. The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: Evidence from 210 patients with stroke. Brain 2009; 132: 3401-10.

Leisenring W, Alono T, Pepe MS. Comparisons of predictive values of binary medical diagnostic tests for paired designs. Biometrics 2000; 56: 345-51.

Lemm S, Blankertz B, Dickhaus T, Müller K-R. Introduction to machine learning for brain imaging. Neuroimage 2011; 56: 387-99.

Lendrem W, Lincoln NB. Spontaneous recovery of language in patients with aphasia between 4 and 34 weeks after stroke. J Neurol Neurosurg Psychiatry 1985; 48: 743-8.

Leonard MK, Cai R, Babiak MC, Ren A, Chang EF. The peri-sylvian cortical network underlying single word repetition revealed by electrocortical stimulation and direct neural recordings. Brain Lang 2016. 10.1016/j.bandl.2016.06.001

Leonard MK, Chang EF. Dynamic speech representations in the human temporal lobe. Trends Cogn Sci 2014; 18: 472-9.

Levelt WJ, Roelofs A, Meyer AS. A theory of lexical access in speech production. Behav Brain Sci 1999; 22: 1-38; discussion 38-75.

Lichtheim L. On aphasia. Brain 1985; 7: 433-84.

Liebenthal E, Sabri M, Beardsley SA, Mangalathu-Arumana J, Desai A. Neural dynamics of phonological processing in the dorsal auditory stream. J Neurosci 2013; 33: 15414-24.

Liew S-L, Santarnecchi E, Buch ER, Cohen LG. Non-invasive brain stimulation in neurorehabilitation: Local and distant effects for motor recovery. Front Hum Neurosci 2014; 8: 378.

Linden DEJ. The working memory networks of the human brain. Neuroscientist 2007; 13: 257-67.

Llano DA, Small SL. Pharmacotherapy for aphasia. In: Hickok G, Small SL, editors. Neurobiology of language. San Diego: Academic Press; 2016. p. 1067-83.

Lo EH, Dalkara T, Moskowitz MA. Mechanisms, challenges and opportunities in stroke. Nat Rev Neurosci 2003; 4: 399-415.

Logothetis NK. What we can do and what we cannot do with fMRI. Nature 2008; 453:869-78.

Logothetis NK, Augath M, Murayama Y, Rauch A, Sultan F, Goense J, et al. The effects of electrical microstimulation on cortical signal propagation. Nat Neurosci 2010; 13:1283-91.

Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. Nature 2001; 412: 150-7.

Logothetis NK, Wandell BA. Interpreting the BOLD signal. Annu Rev Physiol 2004; 66: 735-69.

Lorca-Puls DL, Gajardo-Vidal A, White J, Seghier ML, Leff AP, Green DW, et al. The impact of sample size on the reproducibility of voxel-based lesion-deficit mappings. Neuropsychologia 2018; 115: 101-11.

Lukic S, Barbieri E, Wang X, Caplan D, Kiran S, Rapp B, et al. Right hemisphere grey matter volume and language functions in stroke aphasia. Neural Plast 2017; 2017: 5601509.

Maas E, Gutiérrez K, Ballard KJ. Phonological encoding in apraxia of speech and aphasia. Aphasiology 2014; 28: 25-48.
Mah Y-H, Husain M, Rees G, Nachev P. Human brain lesion-deficit inference remapped. Brain 2014; 137: 2522-31.
Majerus S. Language repetition and short-term memory: An integrative framework. Front Hum Neurosci 2013; 7: 357.
Majerus S, Linden MVd, Mulder L, Meulemans T, Peters F. Verbal short-term memory reflects the sublexical organization of the phonological language network: Evidence from an incidental phonotactic learning paradigm. J Mem Lang 2004; 51: 297-306.
Makris N, Kennedy DN, McInerney S, Sorensen AG, Wang R, Caviness JVS, et al. Segmentation of subcomponents within the superior longitudinal fascicle in humans: A quantitative, in vivo, DT-MRI study. Cereb Cortex 2005; 15: 854-69.
Maldonado IL, Moritz-Gasser S, Duffau H. Does the left superior longitudinal fascicle subserve language semantics? A brain electrostimulation study. Brain Struct Funct 2011; 216: 263-74.
Malherbe C, Umarova RM, Zavaglia M, Kaller CP, Beume L, Thomalla G, et al. Neural correlates of visuospatial bias in patients with left hemisphere stroke: A causal functional contribution analysis based on game theory. Neuropsychologia 2018; 115: 142-53.
Malins JG, Gumkowski N, Buis B, Molfese P, Rueckl JG, Frost SJ, et al. Dough, tough, cough, rough: A “fast” fMRI localizer of component processes in reading. Neuropsychologia 2016; 91: 394-406.
Mandelli ML, Caverzasi E, Binney RJ, Henry ML, Lobach I, Block N, et al. Frontal white matter tracts sustaining speech production in primary progressive aphasia. J Neurosci 2014; 34: 9754-67.
Manuel AL, Radman N, Mesot D, Chouiter L, Clarke S, Annoni J-M, et al. Inter- and intrahemispheric dissociations in ideomotor apraxia: A large-scale lesion–symptom mapping study in subacute brain-damaged patients. Cereb Cortex 2013; 23: 2781-9.
Marchina S, Zhu Lin L, Norton A, Zipse L, Wan Catherine Y, Schlaug G. Impairment of speech production predicted by lesion load of the left arcuate fasciculus. Stroke 2011; 42: 2251-6.
Margulies DS, Petrides M. Distinct parietal and temporal connectivity profiles of ventrolateral frontal areas involved in language production. J Neurosci 2013; 33: 16846-52.
Marie P. Nouveau cas d'aphasie de Broca sans lésions de la troisième frontale gauche. Bulletins et Mémoires de la Société Médicale des Hôpitaux de Paris 1906; 3: 1180-3.
Markiewicz CJ, Bohland JW. Mapping the cortical representation of speech sounds in a syllable repetition task. Neuroimage 2016; 141: 174-90.
Marshall J. Classification of aphasia: Are there benefits for practice? Aphasiology 2010; 24: 408-12.
Martin M, Dressing A, Bormann T, Schmidt CSM, Kümmerer D, Beume L, et al. Componential network for the recognition of tool-associated actions: Evidence from voxel-based lesion-symptom mapping in acute stroke patients. Cereb Cortex 2017; 27: 4139-52.
Martin N, Saffran EM. Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. Cogn Neuropsychol 1997; 14: 641-82.

Martin N, Saffran EM. Effects of word processing and short-term memory deficits on verbal learning: Evidence from aphasia. Int J Psychol 1999; 34: 339-46.
Martin N, Saffran EM. The relationship of input and output phonological processing: An evaluation of models and evidence to support them. Aphasiology 2002; 16: 107-50.
Martin RC, Lesch MF, Bartha MC. Independence of input and output phonology in word processing and short-term memory. J Mem Lang 1999; 41: 3-29.
Martin RC, Shelton JR, Yaffee LS. Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. J Mem Lang 1994; 33: 83-111.
Martino J, De Witt Hamer PC, Berger MS, Lawton MT, Arnold CM, de Lucas EM, et al. Analysis of the subcomponents and cortical terminations of the perisylvian superior longitudinal fasciculus: A fiber dissection and DTI tractography study. Brain Struct Funct 2013; 218: 105-21.
Mateos-Pérez JM, Dadar M, Lacalle-Aurioles M, Iturria-Medina Y, Zeighami Y, Evans AC. Structural neuroimaging as clinical predictor: A review of machine learning applications. Neuroimage Clin 2018; 20: 506-22.
Matsumoto R, Nair DR, LaPresto E, Najm I, Bingaman W, Shibasaki H, et al. Functional connectivity in the human language system: A cortico-cortical evoked potential study. Brain 2004; 127: 2316-30.
Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, et al. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos Trans R Soc Lond B Biol Sci 2001a; 356: 1293-322.
Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, et al. A four-dimensional probabilistic atlas of the human brain. J Am Med Inform Assoc 2001b; 8: 401-30.
McCarthy R, Warrington EK. A two-route model of speech production: Evidence from aphasia. Brain 1984; 107: 463-85.
McCarthy RA, Warrington EK. Repeating without semantics: Surface dysphasia? Neurocase 2001; 7: 77-87.
McDermott KB, Petersen SE, Watson JM, Ojemann JG. A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. Neuropsychologia 2003; 41: 293-303.
McGettigan C, Warren JE, Eisner F, Marshall CR, Shanmugalingam P, Scott SK. Neural correlates of sublexical processing in phonological working memory. J Cogn Neurosci 2011; 23: 961-77.
Mechelli A, Crinion JT, Long S, Friston KJ, Lambon Ralph MA, Patterson K, et al. Dissociating reading processes on the basis of neuronal interactions. J Cogn Neurosci 2005; 17: 1753-65.
Mechelli A, Price CJ, Friston KJ, Ashburner J. Voxel-based morphometry of the human brain: Methods and applications. Curr Med Imaging Rev 2005; 1: 105-13.
Mei L, Xue G, Lu Z-L, He Q, Zhang M, Wei M, et al. Artificial language training reveals the neural substrates underlying addressed and assembled phonologies. PLoS One 2014; 9: e93548.
Meister IG, Wilson SM, Deblieck C, Wu AD, Iacoboni M. The essential role of premotor cortex in speech perception. Curr Biol 2007; 17: 1692-6.
Melloni M, Billeke P, Baez S, Hesse E, de la Fuente L, Forno G, et al. Your perspective and my benefit: Multiple lesion models of self-other integration strategies during social bargaining. Brain 2016; 139: 1-19.
Menon RS. The great brain versus vein debate. Neuroimage 2012; 62: 970-4.

Mesgarani N, Cheung C, Johnson K, Chang EF. Phonetic feature encoding in human superior temporal gyrus. Science 2014; 343: 1006-10.
Mesulam MM. Large-scale neurocognitive networks and distributed processing for attention, language, and memory. Ann Neurol 1990; 28: 597-613.
Mesulam MM, Rogalski EJ, Wieneke C, Hurley RS, Geula C, Bigio EH, et al. Primary progressive aphasia and the evolving neurology of the language network. Nat Rev Neurol 2014; 10: 554-69.
Mesulam MM, Thompson CK, Weintraub S, Rogalski EJ. The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. Brain 2015; 138: 2423-37.
Meyer AM, Snider SF, Campbell RE, Friedman RB. Phonological short-term memory in logopenic variant primary progressive aphasia and mild Alzheimer's disease. Cortex 2015; 71: 183-9.
Michel F, Andreewsky E. Deep dysphasia: An analog of deep dyslexia in the auditory modality. Brain Lang 1983; 18: 212-23.
Miniussi C, Harris JA, Ruzzoli M. Modelling non-invasive brain stimulation in cognitive neuroscience. Neurosci Biobehav Rev 2013; 37: 1702-12.
Mirman D, Chen Q, Zhang Y, Wang Z, Faseyitan OK, Coslett HB, et al. Neural organization of spoken language revealed by lesion–symptom mapping. Nat Commun 2015; 6: 6762.
Mohr JP, Pessin MS, Finkelstein S, Funkenstein HH, Duncan GW, Davis KR. Broca aphasia: Pathologic and clinical. Neurology 1978; 28: 311-24.
Moro V, Urgesi C, Pernigo S, Lanteri P, Pazzaglia M, Aglioti SM. The neural basis of body form and body action agnosia. Neuron 2008; 60: 235-46.
Morris R, Eccles A, Ryan B, Kneebone II. Prevalence of anxiety in people with aphasia after stroke. Aphasiology 2017; 31: 1410-5.
Moser D, Baker JM, Sanchez CE, Rorden C, Fridriksson J. Temporal order processing of syllables in the left parietal lobe. J Neurosci 2009; 29: 12568-73.
Mugler JP, Brookeman JR. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). Magn Reson Med 1990; 15: 152-7.
Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. Science 2005; 309: 951-4.
Munding D, Dubarry A-S, Alario FX. On the cortical dynamics of word production: A review of the MEG evidence. Lang Cogn Neurosci 2016; 31: 441-62.
Murakami T, Kell CA, Restle J, Ugawa Y, Ziemann U. Left dorsal speech stream components and their contribution to phonological processing. J Neurosci 2015; 35: 1411-22.
Naghavi M, Abajobir AA, Abbafati C, Abbas KM, Abd-Allah F, Abera SF, et al. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017; 390: 1151-210.
Nair DG. About being BOLD. Brain Res Rev 2005; 50: 229-43.
Nakamura K, Hara N, Kouider S, Takayama Y, Hanajima R, Sakai K, et al. Task-guided selection of the dual neural pathways for reading. Neuron 2006; 52: 557-64.
Nathan GS. Phonology. In: Dancygier B, editor. The cambridge handbook of cognitive linguistics. Cambridge: Cambridge University Press; 2017. p. 214-28.
Neggers SFW, Petrov PI, Mandija S, Sommer IEC, van den Berg NAT. Understanding the biophysical effects of transcranial magnetic stimulation on brain tissue: The

bridge between brain stimulation and cognition. In: Bestmann S, editor. Computational neurostimulation. Prog Brain Res: Elsevier; 2015. p. 229-59.
Nickels L, Howard D, Best W. Fractionating the articulatory loop: Dissociations and associations in phonological recoding in aphasia. Brain Lang 1997; 56: 161-82.
Nimmo LM, Roodenrys S. Investigating the phonological similarity effect: Syllable structure and the position of common phonemes. J Mem Lang 2004; 50: 245-58.
Nixon P, Lazarova J, Hodinott-Hill I, Gough P, Passingham R. The inferior frontal gyrus and phonological processing: An investigation using rTMS. J Cogn Neurosci 2004; 16: 289-300.
Niziolek CA, Nagarajan SS, Houde JF. What does motor efference copy represent? Evidence from speech production. J Neurosci 2013; 33: 16110-6.
Noppeney U, Penny WD, Price CJ, Flandin G, Friston KJ. Identification of degenerate neuronal systems based on intersubject variability. Neuroimage 2006; 30: 885-90.
Nozari N, Kittredge AK, Dell GS, Schwartz MF. Naming and repetition in aphasia: Steps, routes, and frequency effects. J Mem Lang 2010; 63: 541-59.
Nummenmaa A, McNab JA, Savadjiev P, Okada Y, Hämäläinen MS, Wang R, et al. Targeting of white matter tracts with transcranial magnetic stimulation. Brain Stimul 2014; 7: 80-4.
O'Brien JT, Thomas A. Vascular dementia. Lancet 2015; 386: 1698-706.
O'Donnell MJ, Chin SL, Rangarajan S, Xavier D, Liu L, Zhang H, et al. Global and regional effects of potentially modifiable risk factors associated with acute stroke in 32 countries (INTERSTROKE): A case-control study. Lancet 2016; 388: 761-75.
Oberhuber M, Hope TMH, Seghier ML, Parker Jones O, Prejawa S, Green DW, et al. Four functionally distinct regions in the left supramarginal gyrus support word processing. Cereb Cortex 2016; 26: 4212-26.
Ogawa S, Lee T-M, Nayak AS, Glynn P. Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields. Magn Reson Med 1990a; 14: 68-78.
Ogawa S, Lee TM, Kay AR, Tank DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proc Natl Acad Sci USA 1990b; 87: 9868.
Okada K, Matchin W, Hickok G. Phonological feature repetition suppression in the left inferior frontal gyrus. J Cogn Neurosci 2018; 30: 1549-57.
Oppenheim GM, Dell GS. Motor movement matters: The flexible abstractness of inner speech. Mem Cognit 2010; 38: 1147-60.
Page MPA, Madge A, Cumming N, Norris DG. Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus. J Mem Lang 2007; 56: 49-64.
Papagno C, Comi A, Riva M, Bizzi A, Vernice M, Casarotti A, et al. Mapping the brain network of the phonological loop. Hum Brain Mapp 2017; 38: 3011-24.
Papoutsi M, de Zwart JA, Jansma JM, Pickering MJ, Bednar JA, Horwitz B. From phonemes to articulatory codes: An fMRI study of the role of broca's area in speech production. Cereb Cortex 2009; 19: 2156-65.
Parker Jones Ō, Prejawa S, Hope T, Oberhuber M, Seghier M, Leff A, et al. Sensory-to-motor integration during auditory repetition: A combined fMRI and lesion study. Front Hum Neurosci 2014; 8: 24.
Parkin BL, Ekhtiari H, Walsh VF. Non-invasive human brain stimulation in cognitive neuroscience: A primer. Neuron 2015; 87: 932-45.
Pascual-Leone A, Bartres-Fazf D, Keenan JP. Transcranial magnetic stimulation: Studying the brain-behaviour relationship by induction of ‘virtual lesions’. Philos Trans R Soc Lond B Biol Sci 1999; 354:1229-38.

Pascual-Leone A, Walsh V, Rothwell J. Transcranial magnetic stimulation in cognitive neuroscience – virtual lesion, chronometry, and functional connectivity. Curr Opin Neurobiol 2000; 10: 232-7.

Patterson K, Shewell C. Speak and spell: Dissociations and word-class effects. In: Coltheart M, Job R, Sartori G, editors. The cognitive neuropsychology of language. Hove: Lawrence Erlbaum Associates; 1987. p. 273-94.

Paulesu E, Frith CD, Frackowiak RSJ. The neural correlates of the verbal component of working memory. Nature 1993; 362: 342-5.

Pedersen PM, Jørgensen HS, Nakayama H, Raaschou HO, Olsen TS. Aphasia in acute stroke: Incidence, determinants, and recovery. Ann Neurol 1995; 38: 659-66.

Pedersen P, Vinter K, Olsen TS. Aphasia after stroke: Type, severity and prognosis. Cerebrovasc Dis 2004; 17: 35-43.

Peeva MG, Guenther FH, Tourville JA, Nieto-Castanon A, Anton J-L, Nazarian B, et al. Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. Neuroimage 2010; 50: 626-38.

Pei X, Leuthardt EC, Gaona CM, Brunner P, Wolpaw JR, Schalk G. Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. Neuroimage 2011; 54: 2960-72.

Pell GS, Roth Y, Zangen A. Modulation of cortical excitability induced by repetitive transcranial magnetic stimulation: Influence of timing and geometrical parameters and underlying mechanisms. Prog Neurobiol 2011; 93: 59-98.

Pendlebury ST, Rothwell PM. Prevalence, incidence, and factors associated with pre-stroke and post-stroke dementia: A systematic review and meta-analysis. Lancet Neurol 2009; 8: 1006-18.

Peramunage D, Blumstein SE, Myers EB, Goldrick M, Baese-Berk M. Phonological neighborhood effects in spoken word production: An fMRI study. J Cogn Neurosci 2011; 23: 593-603.

Perrachione TK, Ghosh SS, Ostrovskaya I, Gabrieli JDE, Kovelman I. Phonological working memory for words and nonwords in cerebral cortex. J Speech Lang Hear Res 2017; 60: 1959-79.

Perrone-Bertolotti M, Rapin L, Lachaux JP, Baciu M, Lœvenbruck H. What is that little voice inside my head? Inner speech phenomenology, its role in cognitive performance, and its relation to self-monitoring. Behav Brain Res 2014; 261: 220-39.

Peschke C, Ziegler W, Eisenberger J, Baumgaertner A. Phonological manipulation between speech perception and production activates a parieto-frontal circuit. Neuroimage 2012; 59: 788-99.

Peterchev AV, Wagner TA, Miranda PC, Nitsche MA, Paulus W, Lisanby SH, et al. Fundamentals of transcranial electric and magnetic stimulation dose: Definition, selection, and reporting practices. Brain Stimul 2012; 5: 435-53.

Petrides M, Pandya DN. Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. PLoS Biol 2009; 7: e1000170.

Phan T, G., Donnan G, A., Wright P, M., Reutens DC. A digital map of middle cerebral artery infarcts associated with middle cerebral artery trunk and branch occlusion. Stroke 2005; 36: 986-91.

Philipose LE, Gottesman RF, Newhart M, Kleinman JT, Herskovits EH, Pawlak MA, et al. Neural regions essential for reading and spelling of words and pseudowords. Ann Neurol 2007; 62: 481-92.

Pickering MJ, Clark A. Getting ahead: forward models and their place in cognitive architecture. Trends Cogn Sci 2014; 18: 451-6.

Pike C, Kritzinger A, Pillay B. Social participation in working-age adults with aphasia: An updated systematic review. Top Stroke Rehabil 2017; 24: 627-39.
Pillay SB, Stengel BC, Humphries C, Book DS, Binder JR. Cerebral localization of impaired phonological retrieval during rhyme judgment. Ann Neurol 2014; 76: 738-46.
Plewes DB, Kucharczyk W. Physics of MRI: A primer. J Magn Reson Imaging 2012; 35: 1038-54.
Polanía R, Nitsche MA, Ruff CC. Studying and modifying brain function with non-invasive brain stimulation. Nat Neurosci 2018; 21: 174-87.
Poldrack RA. The role of fMRI in cognitive neuroscience: Where do we stand? Curr Opin Neurobiol 2008; 18: 223-7.
Poldrack RA, Mumford JA, Nichols, TE, editors. Handbook of functional MRI data analysis. New York: Cambridge University Press; 2011.
Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JDE. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage 1999; 10: 15-35.
Pollock A, St George B, Fenton M, Firkins L. Top 10 research priorities relating to life after stroke – consensus from stroke survivors, caregivers, and health professionals. Int J Stroke 2014; 9: 313-20.
Pooley RA. Fundamental physics of MR imaging. Radiographics 2005; 25: 1087-99.
Popovici V, Chen W, Gallas BD, Hatzis C, Shi W, Samuelson FW, et al. Effect of training-sample size and classification difficulty on the accuracy of genomic predictors. Breast Cancer Res 2010; 12: R5.
Price CJ. The anatomy of language: A review of 100 fMRI studies published in 2009. Ann N Y Acad Sci 2010; 1191: 62-88.
Price CJ. A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage 2012; 62: 816-47.
Price CJ, Friston KJ. Degeneracy and cognitive anatomy. Trends Cogn Sci 2002; 6: 416-21.
Price CJ, Hope TM, Seghier ML. Ten problems and solutions when predicting individual outcome from lesion site after stroke. Neuroimage 2017; 145: 200-8.
Price CJ, Moore CJ, Humphreys GW, Wise RJS. Segregating semantic from phonological processes during reading. J Cogn Neurosci 1997; 9: 727-33.
Price CJ, Mummery CJ, Moore CJ, Frackowiak RSJ, Friston KJ. Delineating necessary and sufficient neural systems with functional imaging studies of neuropsychological patients. J Cogn Neurosci 1999; 11: 371-82.
Price CJ, Ramsden S, Hope TMH, Friston KJ, Seghier ML. Predicting IQ change from brain structure: A cross-validation study. Dev Cogn Neurosci 2013; 5: 172-84.
Price CJ, Seghier ML, Leff AP. Predicting language outcome and recovery after stroke: The PLORAS system. Nat Rev Neurol 2010a; 6:202-10.
Price CJ, Crinion JT, Leff AP, Richardson FM, Schofield TM, Prejawa S, et al. Lesion sites that predict the ability to gesture how an object is used. Arch Ital Biol 2010b; 148: 243-58.
Price CJ, Winterburn D, Giraud AL, Moore CJ, Noppeney U. Cortical localisation of the visual and auditory word form areas: A reconsideration of the evidence. Brain Lang 2003; 86: 272-86.
Pustina D, Avants B, Faseyitan OK, Medaglia JD, Coslett HB. Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. Neuropsychologia 2018; 115: 154-66.

Pustina D, Coslett HB, Ungar L, Faseyitan OK, Medaglia JD, Avants B, et al. Enhanced estimations of post-stroke aphasia severity using stacked multimodal predictions. Hum Brain Mapp 2017; 38: 5603-15.

Radua J, Mataix-Cols D. Voxel-wise meta-analysis of grey matter changes in obsessive–compulsive disorder. Br J Psychiatry 2009; 195: 393-402.

Raizada RDS, Poldrack RA. Selective amplification of stimulus differences during categorical processing of speech. Neuron 2007; 56: 726-40.

Rapcsak SZ, Beeson PM, Henry ML, Leyden A, Kim E, Rising K, et al. Phonological dyslexia and dysgraphia: Cognitive mechanisms and neural substrates. Cortex 2009; 45: 575-91.

Rapcsak SZ, Henry ML, Teague SL, Carnahan SD, Beeson PM. Do dual-route models accurately predict reading and spelling performance in individuals with acquired alexia and agraphia? Neuropsychologia 2007; 45: 2519-24.

Rapp B, Goldrick M. Discreteness and interactivity in spoken word production. Psychol Rev 2000; 107: 460-99.

Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. Nat Neurosci 2009; 12: 718-24.

Ravazzani P, Ruohonen J, Grandori F, Tognola G. Magnetic stimulation of the nervous system: Induced electric field in unbounded, semi-infinite, spherical, and cylindrical media. Ann Biomed Eng 1996; 24: 606-16.

Ravizza SM, Hazeltine E, Ruiz S, Zhu DC. Left TPJ activity in verbal working memory: Implications for storage- and sensory-specific models of short term memory. Neuroimage 2011; 55: 1836-46.

Rehme AK, Volz LJ, Feis D-L, Eickhoff SB, Fink GR, Grefkes C. Individual prediction of chronic motor outcome in the acute post-stroke stage: Behavioral parameters versus functional imaging. Hum Brain Mapp 2015a; 36: 4553-65.

Rehme AK, Volz LJ, Feis DL, Bomilcar-Focke I, Liebig T, Eickhoff SB, et al. Identifying neuroimaging markers of motor disability in acute stroke by machine learning techniques. Cereb Cortex 2015b; 25: 3046-56.

Robson H, Keidel JL, Lambon Ralph MA, Sage K. Revealing and quantifying the impaired phonological analysis underpinning impaired comprehension in Wernicke's aphasia. Neuropsychologia 2012; 50: 276-88.

Roelofs A. The WEAVER model of word-form encoding in speech production. Cognition 1997; 64: 249-84.

Roelofs A. A dorsal-pathway account of aphasic language production: The WEAVER++/ARC model. Cortex 2014; 59: 33-48.

Rogalsky C, Poppa T, Chen K-H, Anderson SW, Damasio H, Love T, et al. Speech repetition as a window on the neurobiology of auditory–motor integration for speech: A voxel-based lesion symptom mapping study. Neuropsychologia 2015; 71: 18-27.

Roh JH, Kim JH, Park K-W, Pyun S-B, Kim S-E, Hwang YM. Wernicke's aphasia caused by Broca's area infarction: A PET follow-up study. Alzheimers Dement 2009; 5: P201-2.

Rolheiser T, Stamatakis EA, Tyler LK. Dynamic processing in the human language system: Synergy between the arcuate fascicle and extreme capsule. J Neurosci 2011; 31: 16949-57.

Rolston JD, Chang EF. Critical language areas show increased functional connectivity in human cortex. Cereb Cortex 2018; 28: 4161-8.

Romero L, Walsh V, Papagno C. The neural correlates of phonological short-term memory: A repetitive transcranial magnetic stimulation study. J Cogn Neurosci 2006; 18: 1147-55.

Rondina JM, Filippone M, Girolami M, Ward NS. Decoding post-stroke motor function from structural brain imaging. Neuroimage Clin 2016; 12: 372-80.

Rorden C, Fridriksson J, Karnath H-O. An evaluation of traditional and novel tools for lesion behavior mapping. Neuroimage 2009; 44: 1355-62.

Rorden C, Karnath H-O. Using human brain lesions to infer function: A relic from a past era in the fMRI age? Nat Rev Neurosci 2004; 5: 813-9.

Rorden C, Karnath H-O, Bonilha L. Improving lesion-symptom mapping. J Cogn Neurosci 2007; 19: 1081-8.

Rosen HJ, Gorno-Tempini ML, Goldman WP, Perry RJ, Schuff N, Weiner M, et al. Patterns of brain atrophy in frontotemporal dementia and semantic dementia. Neurology 2002; 58: 198-208.

Rossi S, Hallett M, Rossini PM, Pascual-Leone A. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. Clin Neurophysiol 2009; 120: 2008-39.

Rossini PM, Burke D, Chen R, Cohen LG, Daskalakis Z, Di Iorio R, et al. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: Basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. committee. Clin Neurophysiol 2015; 126: 1071-107.

Roth GA, Forouzanfar MH, Moran AE, Barber R, Nguyen G, Feigin VL, et al. Demographic and epidemiologic drivers of global cardiovascular mortality. N Engl J Med 2015; 372: 1333-41.

Roux F-E, Durand J-B, Jucla M, Réhault E, Reddy M, Démonet J-F. Segregation of lexical and sub-lexical reading processes in the left perisylvian cortex. PLoS One 2012; 7: e50665.

Ruchkin DS, Grafman J, Cameron K, Berndt RS. Working memory retention systems: A state of activated long-term memory. Behav Brain Sci 2003; 26: 709-28.

Rudrauf D, Mehta S, Bruss J, Tranel D, Damasio H, Grabowski TJ. Thresholding lesion overlap difference maps: Application to category-related naming and recognition deficits. Neuroimage 2008; 41: 970-84.

Ruschel M, Knösche TR, Friederici AD, Turner R, Geyer S, Anwander A. Connectivity architecture and subdivision of the human inferior parietal cortex revealed by diffusion MRI. Cereb Cortex 2014; 24: 2436-48.

Sandrini M, Umiltà C, Rusconi E. The use of transcranial magnetic stimulation in cognitive neuroscience: A new synthesis of methodological issues. Neurosci Biobehav Rev 2011; 35: 516-36.

Sato M, Schwartz J-L, Perrier P. Phonemic auditory and somatosensory goals in speech production. Lang Cogn Neurosci 2014; 29: 41-3.

Sato M, Tremblay P, Gracco VL. A mediating role of the premotor cortex in phoneme segmentation. Brain Lang 2009; 111: 1-7.

Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry M-S, et al. Ventral and dorsal pathways for language. Proc Natl Acad Sci USA 2008; 105: 18035-40.

Saur D, Lange R, Baumgaertner A, Schraknepper V, Willmes K, Rijntjes M, et al. Dynamics of language reorganization after stroke. Brain 2006; 129: 1371-84.

Savill NJ, Cornelissen P, Pahor A, Jefferies E. rTMS evidence for a dissociation in short-term memory for spoken words and nonwords. Cortex 2018. 10.1016/j.cortex.2018.07.021

Saygin AP. Superior temporal and premotor brain areas necessary for biological motion perception. Brain 2007; 130: 2452-61.
Schmitz J, Díaz B, Fernández Rubio K, Sebastian-Galles N. Exploring the relationship between speech perception and production across phonological processes, language familiarity, and sensory modalities. Lang Cogn Neurosci 2018; 33: 527-46.
Schwartz MF. Theoretical analysis of word production deficits in adult aphasia. Philos Trans R Soc Lond B Biol Sci 2014; 369: 20120390.
Schwartz MF, Dell GS. Word production from the perspective of speech errors in aphasia. In: Hickok G, Small SL, editors. Neurobiology of language. San Diego: Academic Press; 2016. p. 701-15.
Schwartz MF, Faseyitan O, Kim J, Coslett HB. The dorsal stream contribution to phonological retrieval in object naming. Brain 2012; 135: 3799-814.
Schweppe J, Grice M, Rummer R. What models of verbal working memory can learn from phonological theory: Decomposing the phonological similarity effect. J Mem Lang 2011; 64: 256-69.
Seghier ML, Bagdasaryan J, Jung DE, Price CJ. The importance of premotor cortex for supporting speech production after left capsular-putaminal damage. J Neurosci 2014; 34: 14338-48.
Seghier ML, Friston KJ, Price CJ. Detecting subject-specific activations using fuzzy clustering. Neuroimage 2007; 36: 594-605.
Seghier ML, Lazeyras F, Pegna AJ, Annoni J-M, Zimine I, Mayer E, et al. Variability of fMRI activation during a phonological and semantic language task in healthy subjects. Hum Brain Mapp 2004; 23: 140-55.
Seghier ML, Ramlackhansingh A, Crinion J, Leff AP, Price CJ. Lesion identification using unified segmentation-normalisation models and fuzzy clustering. Neuroimage 2008a; 41: 1253-66.
Seghier ML, Lee HL, Schofield T, Ellis CL, Price CJ. Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. Neuroimage 2008b; 42: 1226-36.
Seghier ML, Neufeld NH, Zeidman P, Leff AP, Mechelli A, Nagendran A, et al. Reading without the left ventral occipito-temporal cortex. Neuropsychologia 2012; 50: 3621-35.
Seghier ML, Patel E, Prejawa S, Ramsden S, Selmer A, Lim L, et al. The PLORAS Database: A data repository for predicting language outcome and recovery after stroke. Neuroimage 2016; 124: 1208-12.
Seghier ML, Price CJ. Dissociating functional brain networks by decoding the between-subject variability. Neuroimage 2009; 45: 349-59.
Seghier ML, Price CJ. Reading aloud boosts connectivity through the putamen. Cereb Cortex 2010; 20: 570-82.
Seghier ML, Price CJ. Visualising inter-subject variability in fMRI using threshold-weighted overlap maps. Sci Rep 2016; 6: 20170.
Seghier ML, Price CJ. Interpreting and utilising intersubject variability in brain function. Trends Cogn Sci 2018; 22: 517-30.
Seghier ML, Zeidman P, Neufeld N, Leff AP, Price CJ. Identifying abnormal connectivity in patients using Dynamic Causal Modelling of fMRI responses. Front Syst Neurosci 2010; 4: 142.
Seidenberg MS. Computational models of reading. In: Spivey M, McRae K, Joanisse M, editors. The cambridge handbook of psycholinguistics. Cambridge: Cambridge University Press; 2012. p. 186-203.

Selnes OA, Niccum N, Knopman DS, Rubens AB. Recovery of single word comprehension: CT-scan correlates. Brain Lang 1984; 21: 72-84.
Selnes OA, van Zijl PCM, Barker PB, Hillis AE, Mori S. MR diffusion tensor imaging documented arcuate fasciculus lesion in a patient with normal repetition performance. Aphasiology 2002; 16: 897-902.
Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, et al. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science 1995; 268: 889-93.
Service E. Phonology, working memory, and foreign-language learning. Q J Exp Psychol A 1992; 45: 21-50.
Shallice T, editor. From neuropsychology to mental structure. Cambridge: Cambridge University Press; 1988.
Shallice T, Rumiati RI, Zadini A. The selective impairment of the phonological output buffer. Cogn Neuropsychol 2000; 17: 517-46.
Shehata GA, El Mistikawi T, Risha ASK, Hassan HS. The effect of aphasia upon personality traits, depression and anxiety among stroke patients. J Affect Disord 2015; 172: 312-4.
Sheth SA, Nemoto M, Guiou MW, Walker MA, Toga AW. Spatiotemporal evolution of functional hemodynamic changes and their relationship to neuronal activity. J Cereb Blood Flow Metab 2005; 25: 830-41.
Shimotake A, Matsumoto R, Ueno T, Kunieda T, Saito S, Hoffman P, et al. Direct exploration of the role of the ventral anterior temporal lobe in semantic memory: Cortical stimulation and local field potential evidence from subdural grid electrodes. Cereb Cortex 2015; 25: 3802-17.
Shmueli G. To explain or to predict? Statist Sci 2010; 25: 289-310.
Shuster LI, Lemieux SK. An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. Brain Lang 2005; 93: 20-31.
Siebner HR, Hartwigsen G, Kassuba T, Rothwell JC. How does transcranial magnetic stimulation modify neuronal activity in the brain? Implications for studies of cognition. Cortex 2009; 45: 1035-42.
Siegel JS, Ramsey LE, Snyder AZ, Metcalf NV, Chacko RV, Weinberger K, et al. Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke. Proc Natl Acad Sci USA 2016; 113: E4367-76.
Silveri MC, Cappa A. Segregation of the neural correlates of language and phonological short-term memory. Cortex 2003; 39: 913-25.
Simard F, Monetta L, Nagano-Saito A, Monchi O. A new lexical card-sorting task for studying fronto-striatal contribution to processing language rules. Brain Lang 2013; 125: 295-306.
Skipper-Kallal LM, Lacey EH, Xing S, Turkeltaub PE. Functional activation independently contributes to naming ability and relates to lesion site in post-stroke aphasia. Hum Brain Mapp 2017a; 38: 2051-66.
Skipper-Kallal LM, Lacey EH, Xing S, Turkeltaub PE. Right hemisphere remapping of naming functions depends on lesion size and location in poststroke aphasia. Neural Plast 2017b; 2017: 8740353.
Skipper JI, Devlin JT, Lametti DR. The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. Brain Lang 2017; 164: 77-105.
Sliwinska MW, James A, Devlin JT. Inferior parietal lobule contributions to visual word recognition. J Cogn Neurosci 2015; 27: 593-604.

Smith DV, Clithero JA, Rorden C, Karnath H-O. Decoding the anatomical network of spatial attention. Proc Natl Acad Sci USA 2013; 110: 1518-23.
Smith EE, Jonides J, Marshuetz C, Koeppe RA. Components of verbal working memory: Evidence from neuroimaging. Proc Natl Acad Sci USA 1998; 95: 876-82.
Specht K, Zahn R, Willmes K, Weis S, Holtel C, Krause BJ, et al. Joint independent component analysis of structural and functional images reveals complex patterns of functional reorganisation in stroke aphasia. Neuroimage 2009; 47: 2057-63.
Sperber C, Karnath H-O. Topography of acute stroke in a sample of 439 right brain damaged patients. Neuroimage Clin 2016; 10: 124-8.
Sperber C, Karnath H-O. Impact of correction factors in human brain lesion-behavior inference. Hum Brain Mapp 2017; 38: 1692-701.
Sperber C, Karnath H-O. On the validity of lesion-behaviour mapping methods. Neuropsychologia 2018; 115: 17-24.
Sperber C, Wiesen D, Karnath H-O. An empirical evaluation of multivariate lesion behaviour mapping using support vector regression. Hum Brain Mapp 2018. 10.1002/hbm.24476
Springer L, Mantey S. The comprehensive aphasia test: A review. Commentary on Howard, Swinburn, and Porter, “Putting the CAT out: What the comprehensive aphasia test has to offer”. Aphasiology 2010; 24: 75-8.
Stahl B, Mohr B, Büscher V, Dreyer FR, Lucchese G, Pulvermüller F. Efficacy of intensive aphasia therapy in patients with chronic stroke: A randomised controlled trial. J Neurol Neurosurg Psychiatry 2018; 89: 586-92.
Stephan KE, Schlagenhauf F, Huys QJM, Raman S, Aponte EA, Brodersen KH, et al. Computational neuroimaging strategies for single patient predictions. Neuroimage 2017; 145: 180-99.
Stinear CM, Barber PA, Petoe M, Anwar S, Byblow WD. The PREP algorithm predicts potential for upper limb recovery after stroke. Brain 2012; 135: 2527-35.
Stock C, Hielscher T. DTComPair: Comparison of binary diagnostic tests in a paired study design. R package version 1.0.3, 2014. https://cran.r-project.org/package=DTComPair
Stoeckel MC, Wittsack H-J, Meisel S, Seitz RJ. Pattern of cortex and white matter involvement in severe middle cerebral artery ischemia. J Neuroimaging 2007; 17: 131-40.
Swinburn K, Porter G, Howard D. Comprehensive aphasia test. Hove: Psychology Press; 2004.
Szaflarski JP, Binder JR, Possing ET, McKiernan KA, Ward BD, Hammeke TA. Language lateralization in left-handed and ambidextrous people: fMRI data. Neurology 2002; 59: 238-44.
Takata N, Sugiura Y, Yoshida K, Koizumi M, Hiroshi N, Honda K, et al. Optogenetic astrocyte activation evokes BOLD fMRI response with oxygen consumption without neuronal activity modulation. Glia 2018; 66: 2013-23.
Tanaka H, Toyonaga T, Hashimoto H. Functional and occupational characteristics predictive of a return to work within 18 months after stroke in Japan: Implications for rehabilitation. Int Arch Occup Environ Health 2014; 87: 445-53.
Tate MC, Herbet G, Moritz-Gasser S, Tate JE, Duffau H. Probabilistic map of critical functional regions of the human cerebral cortex: Broca’s area revisited. Brain 2014; 137: 2773-82.
Taylor JSH, Rastle K, Davis MH. Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. Psychol Bull 2013; 139: 766-91.

Teichmann M, Darcy I, Bachoud-Lévi A-C, Dupoux E. The role of the striatum in phonological processing: Evidence from early stages of Huntington's disease. Cortex 2009; 45: 839-49.
Tettamanti M, Moro A, Messa C, Moresco RM, Rizzo G, Carpinelli A, et al. Basal ganglia and language: Phonology modulates dopaminergic release. Neuroreport 2005; 16: 397-401.
Teufel C, Fletcher PC. The promises and pitfalls of applying computational models to neurological and psychiatric disorders. Brain 2016; 139: 2600-8.
Thielscher A, Kammer T. Linking physics with physiology in TMS: A sphere field model to determine the cortical stimulation site in TMS. Neuroimage 2002; 17: 1117-30.
Tian X, Zarate JM, Poeppel D. Mental imagery of speech implicates two mechanisms of perceptual reactivation. Cortex 2016; 77: 1-12.
Tippett DC, Hillis AE. Vascular aphasia syndromes. In: Hickok G, Small SL, editors. Neurobiology of language. San Diego: Academic Press; 2016. p. 913-22.
Tippett DC, Hillis AE. Where are aphasia theory and management "headed"? F1000Res 2017; 6.
Tochadse M, Halai AD, Lambon Ralph MA, Abel S. Unification of behavioural, computational and neural accounts of word production errors in post-stroke aphasia. Neuroimage Clin 2018; 18: 952-62.
Tourville JA, Guenther FH. The DIVA model: A neural theory of speech acquisition and production. Lang Cogn Process 2011; 26: 952-81.
Tree JJ, Kay J. Phonological dyslexia and phonological impairment: An exception to the rule? Neuropsychologia 2006; 44: 2861-73.
Tremblay P, Dick AS. Broca and Wernicke are dead, or moving past the classic model of language neurobiology. Brain Lang 2016; 162: 60-71.
Tsouli S, Kyritsis AP, Tsagalis G, Virvidaki E, Vemmos KN. Significance of aphasia after first-ever acute stroke: Impact on early and late outcomes. Neuroepidemiology 2009; 33: 96-102.
Twomey T, Waters D, Price CJ, Kherif F, Woll B, MacSweeney M. Identification of the regions involved in phonological assembly using a novel paradigm. Brain Lang 2015; 150: 45-53.
Tyler LK, Marslen-Wilson W, Stamatakis EA. Dissociating neuro-cognitive component processes: Voxel-based correlational methodology. Neuropsychologia 2005; 43: 771-8.
Ueno T, Saito S, Rogers TT, Lambon Ralph MA. Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. Neuron 2011; 72: 385-96.
Uğurbil K, Toth L, Kim D-S. How accurate is magnetic resonance imaging of brain function? Trends Neurosci 2003; 26: 108-14.
Uludağ K, Blinder P. Linking brain vascular physiology to hemodynamic response in ultra-high field MRI. Neuroimage 2018; 168: 279-95.
United Nations, Department of Economic and Social Affairs, Population Division. World Population Prospects: The 2017 Revision. Volume I: Comprehensive Tables (ST/ESA/SER.A/399).
Vaden KI, Piquado T, Hickok G. Sublexical properties of spoken words modulate activity in Broca's area but not superior temporal cortex: Implications for models of speech recognition. J Cogn Neurosci 2011; 23: 2665-74.
Valero-Cabré A, Amengual JL, Stengel C, Pascual-Leone A, Coubard OA. Transcranial magnetic stimulation in basic and clinical neuroscience: A comprehensive review

of fundamental principles and novel insights. Neurosci Biobehav Rev 2017; 83: 381-404.
Vallar G, Di Betta AM, Silveri MC. The phonological short-term store-rehearsal system: Patterns of impairment and neural correlates. Neuropsychologia 1997; 35: 795-812.
van der Flier WM, Skoog I, Schneider JA, Pantoni L, Mok V, Chen CLH, et al. Vascular cognitive impairment. Nat Rev Dis Primers 2018; 4: 18003.
Varoquaux G. Cross-validation failure: Small sample sizes lead to large error bars. Neuroimage 2018; 180: 68-77.
Varoquaux G, Raamana PR, Engemann DA, Hoyos-Idrobo A, Schwartz Y, Thirion B. Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. Neuroimage 2017; 145: 166-79.
Verdon V, Schwartz S, Lovblad K-O, Hauert C-A, Vuilleumier P. Neuroanatomy of hemispatial neglect and its functional components: A study using voxel-based lesion-symptom mapping. Brain 2010; 133: 880-94.
Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, et al. Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. Neuroimage 2006; 30: 1414-32.
Vignolo LA, Boccardi E, Caverni L. Unexpected CT-scan findings in global aphasia. Cortex 1986; 22: 55-69.
Vitevitch MS, Luce PA. When words compete: Levels of processing in perception of spoken words. Psychol Sci 1998; 9: 325-9.
Vitevitch MS, Luce PA. Probabilistic phonotactics and neighborhood activation in spoken word recognition. J Mem Lang 1999; 40: 374-408.
Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017; 390: 1211-59.
Wagner T, Rushmore J, Eden U, Valero-Cabre A. Biophysical foundations underlying TMS: Setting the stage for an effective use of neurostimulation in the cognitive neurosciences. Cortex 2009; 45: 1025-34.
Walsh V, Cowey A. Transcranial magnetic stimulation and cognitive neuroscience. Nat Rev Neurosci 2000; 1:73-9.
Walsh V, Rushworth M. A primer of magnetic stimulation as a tool for neuropsychology. Neuropsychologia 1999; 37: 125-35.
Wang J, Marchina S, Norton A, Wan C, Schlaug G. Predicting speech fluency and naming abilities in aphasic patients. Front Hum Neurosci 2013; 7: 831.
Wang X, Pathak S, Stefaneanu L, Yeh F-C, Li S, Fernandez-Miranda JC. Subcomponents and connectivity of the superior longitudinal fasciculus in the human brain. Brain Struct Funct 2016; 221: 2075-92.
Warren JE, Wise RJS, Warren JD. Sounds do-able: Auditory–motor transformations and the posterior temporal plane. Trends Neurosci 2005; 28: 636-43.
Watila MM, Balarabe SA. Factors predicting post-stroke aphasia recovery. J Neurol Sci 2015; 352: 12-8.
Wernicke C. Der aphasische symptomencomplex: Eine psychologische studie auf anatomischer basis. Breslau: Max Cohn & Weigert; 1874.
Whitford TJ, Jack BN, Pearson D, Griffiths O, Luque D, Harris AWF, et al. Neurophysiological evidence of efference copies to inner speech. eLife 2017; 6: e28197.

Wilke M, de Haan B, Juenger H, Karnath H-O. Manual, semi-automated, and automated delineation of chronic brain lesions: A comparison of methods. Neuroimage 2011; 56: 2038-46.

Willmes K, Poeck K. To what extent can aphasic syndromes be localized? Brain 1993; 116: 1527-40.

Wilson SM, Galantucci S, Tartaglia MC, Rising K, Patterson DK, Henry ML, et al. Syntactic processing depends on dorsal language tracts. Neuron 2011; 72: 397-403.

Wilson SA, Thickbroom GW, Mastaglia FL. Transcranial magnetic stimulation mapping of the motor cortex in normal subjects: The representation of two intrinsic hand muscles. J Neurol Sci 1993; 118: 134-44.

Wimmer H, Schurz M, Sturm D, Richlan F, Klackl J, Kronbichler M, et al. A dual-route perspective on poor reading in a regular orthography: An fMRI study. Cortex 2010; 46: 1284-98.

Winder AT, Echagarruga C, Zhang Q, Drew PJ. Weak correlations between hemodynamic signals and ongoing neural activity during the resting state. Nat Neurosci 2017; 20: 1761-9.

Wolpert DM, Ghahramani Z. Computational principles of movement neuroscience. Nat Neurosci 2000; 3: 1212-7.

Woollams AM, Madrid G, Lambon Ralph MA. Using neurostimulation to understand the impact of pre-morbid individual differences on post-lesion outcomes. Proc Natl Acad Sci USA 2017; 114:12279-84.

Worrall L, Sherratt S, Rogers P, Howe T, Hersh D, Ferguson A, et al. What people with aphasia want: Their goals according to the ICF. Aphasiology 2011; 25: 309-22.

Wray F, Clarke D. Longer-term needs of stroke survivors with communication difficulties living in the community: A systematic review and thematic synthesis of qualitative studies. BMJ Open 2017; 7: e017944.

Xiang H-D, Fonteijn HM, Norris DG, Hagoort P. Topographical functional connectivity pattern in the perisylvian language networks. Cereb Cortex 2010; 20: 549-60.

Xing S, Lacey EH, Skipper-Kallal LM, Jiang X, Harris-Love ML, Zeng J, et al. Right hemisphere grey matter structure and language outcomes in chronic left hemisphere stroke. Brain 2016; 139: 227-41.

Xu T, Jha A, Nachev P. The dimensionalities of lesion-deficit mapping. Neuropsychologia 2018; 115: 134-41.

Yagata SA, Yen M, McCarron A, Bautista A, Lamair-Orosco G, Wilson SM. Rapid recovery from aphasia after infarction of Wernicke's area. Aphasiology 2017; 31: 951-80.

Yagmurlu K, Middlebrooks EH, Tanriover N, Rhoton AL Jr. Fiber tracts of the dorsal language stream in the human brain. J Neurosurg 2016; 124: 1396-405.

Ylinen S, Nora A, Leminen A, Hakala T, Huotilainen M, Shtyrov Y, et al. Two distinct auditory-motor circuits for monitoring speech production as revealed by content-specific suppression of auditory cortex. Cereb Cortex 2015; 25: 1576-86.

Yoo S, Chung J-Y, Jeon H-A, Lee K-M, Kim Y-B, Cho Z-H. Dual routes for verbal repetition: Articulation-based and acoustic–phonetic codes for pseudoword and word repetition, respectively. Brain Lang 2012; 122: 1-10.

Yourganov G, Fridriksson J, Rorden C. Estimating the statistical significance of spatial maps for multivariate lesion-symptom analysis. Cortex 2018; 108: 276-8.

Yourganov G, Fridriksson J, Rorden C, Gleichgerrcht E, Bonilha L. Multivariate connectome-based symptom mapping in post-stroke patients: Networks supporting language and speech. J Neurosci 2016; 36: 6668-79.

Yourganov G, Smith KG, Fridriksson J, Rorden C. Predicting aphasia type from brain damage measured with structural MRI. Cortex 2015; 73: 203-15.

Zhang Y, Kimberg DY, Coslett HB, Schwartz MF, Wang Z. Multivariate lesion-symptom mapping using support vector regression. Hum Brain Mapp 2014; 35: 5861-76.

Ziegler W, Ackermann H, Kappes J. From phonology to articulation: A neurophonetic view. In: Guendouzi J, Loncke F, Williams MJ, editors. The handbook of psycholinguistic and cognitive processes: Perspectives in communication disorders. Hove: Psychology Press; 2010. p. 327-43.

Ziegler W, Aichert I, Staiger A. Apraxia of speech: Concepts and controversies. J Speech Lang Hear Res 2012; 55: S1485-501.

Appendix 1 (Chapter 2)

1 Visual phonological decision task

No.	Session	Block	Stimuli		Sound same?
					1 = Yes; 0 = No
1	practice	0	ceiling	sealing	1
2	practice	0	forest	frost	0
3	practice	0	border	bother	0
4	practice	0	petal	pebble	0
5	practice	0	tyre	tire	1
6	practice	0	hour	our	1
7	practice	0	circle	circus	0
8	practice	0	berry	bury	1
9	practice	0	sausage	sauces	0
10	practice	0	cereal	serial	1
11	practice	0	pillow	plough	0
12	practice	0	allowed	aloud	1
1	test	1	ask	axe	0
2	test	1	heel	heal	1
3	test	1	sail	sale	1
4	test	1	pain	pane	1
5	test	1	honey	hunter	0
6	test	1	doe	dough	1
7	test	1	blood	blush	0
8	test	1	void	volt	0
9	test	1	sweet	suite	1
10	test	1	way	weigh	1
11	test	1	sew	sow	1
12	test	1	jury	duty	0
13	test	1	egg	edge	0
14	test	1	him	hymn	1
15	test	1	shot	shop	0
16	test	1	toe	tow	1
17	test	1	pound	pounce	0
18	test	1	pint	pine	0
19	test	1	chant	chart	0
20	test	1	fair	fare	1
21	test	1	links	lynx	1
22	test	1	mail	mile	0
23	test	1	filter	fillet	0
24	test	1	ware	wear	1
25	test	1	owl	old	0
26	test	1	ate	eight	1
27	test	1	grass	grace	0
28	test	1	ewe	you	1
29	test	1	great	grate	1
30	test	1	road	rode	1
31	test	1	map	mop	0

32	test	1	weed	weir	0
33	test	2	court	corpse	0
34	test	2	foil	fail	0
35	test	2	fall	fault	0
36	test	2	seed	soot	0
37	test	2	four	for	1
38	test	2	won	one	1
39	test	2	beet	beat	1
40	test	2	wrap	rap	1
41	test	2	diary	dairy	0
42	test	2	brake	break	1
43	test	2	shake	shock	0
44	test	2	hear	here	1
45	test	2	bear	bare	1
46	test	2	knows	nose	1
47	test	2	mist	moist	0
48	test	2	mall	mole	0
49	test	2	pray	prey	1
50	test	2	liar	lair	0
51	test	2	navel	novel	0
52	test	2	moose	mouse	0
53	test	2	gown	gone	0
54	test	2	rain	reign	1
55	test	2	boil	bowl	0
56	test	2	ail	ale	1
57	test	2	foul	fowl	1
58	test	2	sight	cite	1
59	test	2	would	wood	1
60	test	2	lad	lid	0
61	test	2	sole	soul	1
62	test	2	groan	grow	0
63	test	2	beach	beech	1
64	test	2	wheel	well	0

2 Visual semantic decision task

No.	Session	Block	Stimuli		Related meaning?
					1 = Yes; 0 = No
1	practice	0	thin	open	0
2	practice	0	bar	cave	0
3	practice	0	door	window	1
4	practice	0	gold	silver	1
5	practice	0	never	now	1
6	practice	0	skirt	hero	0
7	practice	0	lessen	few	1
8	practice	0	hell	lady	0
9	practice	0	many	some	1
10	practice	0	clam	oyster	1
11	practice	0	coffee	them	0
12	practice	0	cheat	factor	0
1	test	1	best	first	1
2	test	1	pull	both	0
3	test	1	cup	sell	0
4	test	1	elm	oak	1
5	test	1	flask	jar	1
6	test	1	haul	soon	0
7	test	1	duke	king	1
8	test	1	wolf	drain	0
9	test	1	slow	quick	1
10	test	1	five	two	1
11	test	1	jam	real	0
12	test	1	lace	sole	1
13	test	1	chain	side	0
14	test	1	bean	lack	0
15	test	1	frog	rest	0
16	test	1	how	why	1
17	test	1	lift	rare	0
18	test	1	fork	spoon	1
19	test	1	inch	mile	1
20	test	1	leaf	stem	1
21	test	1	beak	bill	1
22	test	1	blind	tear	0
23	test	1	wine	stool	0
24	test	1	toy	once	0
25	test	1	pound	dime	1
26	test	1	fleet	pig	0
27	test	1	tar	grace	0
28	test	1	ice	myth	0
29	test	1	fate	wish	1
30	test	1	doe	fawn	1
31	test	1	draw	paint	1
32	test	1	won	scent	0

33	test	2	shape	will	0
34	test	2	oath	guy	0
35	test	2	dye	thing	0
36	test	2	isle	part	0
37	test	2	beer	told	0
38	test	2	thick	age	0
39	test	2	seat	egg	0
40	test	2	cage	jail	1
41	test	2	lake	sea	1
42	test	2	hunt	juice	0
43	test	2	dirt	dust	1
44	test	2	jump	strut	1
45	test	2	gain	care	0
46	test	2	prune	bleak	0
47	test	2	straw	west	0
48	test	2	deck	sail	1
49	test	2	brass	steel	1
50	test	2	ear	staff	0
51	test	2	light	dark	1
52	test	2	feel	mood	1
53	test	2	boy	man	1
54	test	2	dew	dawn	1
55	test	2	toad	joy	0
56	test	2	debt	poor	1
57	test	2	hay	neat	0
58	test	2	went	go	1
59	test	2	chop	slice	1
60	test	2	bush	swim	0
61	test	2	dog	cat	1
62	test	2	crane	duck	1
63	test	2	ash	flame	1
64	test	2	tent	gasp	0