LANGUAGE LATERALITY IN IMPAIRED AND TYPICAL LANGUAGE FUNCTIONING



Abigail Bradshaw

St Anne's college

University of Oxford

Doctor of Philosophy in Experimental Psychology

Acknowledgements

It has been my absolute pleasure to work alongside some wonderful people throughout my PhD, to whom I am greatly indebted for their wisdom and encouragement.

First of all I would like to thank my supervisors Prof. Dorothy Bishop and Dr. Zoe Woodhead. Dorothy has been an incredible role model, from her expert knowledge and industrious work ethic to her passion for research integrity and reproducible science. I feel honoured to have had the opportunity to conduct research with such a legendary researcher. Zoe has been a constant encourager, and has been instrumental in enabling me to thrive throughout my PhD. She has always sought to further my development as a researcher, and I have learnt so much from working with her.

My thanks also go to the whole OSCCI team who have been such a supportive and friendly group within which to work. Special thanks go to Paul Thompson for his time and patience in helping my understanding of statistics and for instilling a love of R coding in me. I am also grateful to Leanne Smith for collecting data for me as part of her undergraduate project.

I would further like to thank Prof. Kate Watkins, who has been a mentor to me ever since I arrived in Oxford for my undergraduate degree. Her encouragement from an early stage was a driving force in my decision to pursue a career in academia. I am incredibly thankful for all that she has invested in me.

I am so grateful for the support and love of my family throughout this PhD, from my father, my mother and my sister. You have championed me throughout my

studies, and looked after me when I have needed time out. Thank you for helping me to work hard while reminding me of the importance of rest. I am also thankful to my fiancé Huw, for his encouragement, humour and perspective. Your calm and patient support have kept me grounded throughout this whole process.

Lastly, I submit this thesis with a heart full of praise to the Lord, for his steadfast faithfulness to me through times of joy and times of hardship. Because of his grace I can truthfully say that my worth and pride is not in what I accomplish, but in what he has done for me.

'Whatever you do, work at it with all your heart, as working for the Lord, not for human masters' (Colossians 3:23).

Thesis Abstract

Cerebral lateralisation for language refers to the well-established finding of a bias in language processing towards the left hemisphere. A longstanding question in the field concerns whether such lateralisation is beneficial for language functioning, and in particular whether its disruption may be an aetiological risk factor for disorders of speech and language. The current thesis aimed to take a new approach to this question by considering the importance of within individual variability in lateralisation.

The first part of this thesis considers optimal practice for measurement of lateralisation for language using functional magnetic resonance imaging (fMRI). Chapters two and three present systematic reviews of this literature which discuss the challenges in calculating a laterality index (LI), and in selecting a language activation paradigm respectively. Recommendations are made for optimal laterality measurement with a view to promoting standardisation in protocols across the field.

The second part of this thesis investigates differences in lateralisation and interhemispheric interactions in a group of adults with developmental disorders affecting language. My first experimental study (*Chapter four*) established that this group demonstrated significantly impaired language skills compared to a group of typically developing control adults. The second experiment of my thesis (*Chapter five*) went on to investigate the relationship between language skills and patterns of lateralisation within this same sample. Three language tasks were used with functional transcranial Doppler ultrasound (fTCD) to measure consistency in lateralisation across different language functions. Contrary to

predictions, inconsistency in lateralisation was not associated with poorer language skills across the whole sample. However, the developmental disorder group demonstrated a higher incidence of inconsistent lateralisation, weaker correlations between task LIs, and increased variability in lateralised responses over trials within a phonological task. These findings suggest that inconsistent lateralisation does not directly cause impaired language skills, but nevertheless is more likely to occur in individuals with developmental disorders associated with poorer language functioning.

In a further experiment (*Chapter six*), interhemispheric transfer during word recognition was investigated in adults with dyslexia and typically reading adults using a visual half-field (VHF) paradigm. Mixed evidence for a deficit in interhemispheric transfer in dyslexia was found. Both dyslexic and control participants were able to benefit from simultaneous presentation of identical word stimuli to both visual fields, an effect which is attributed to interhemispheric cooperation. However, dyslexic participants demonstrated poorer performance for left visual field words, suggesting an impairment in the transfer of visual word form information from the right to the (dominant) left hemisphere.

Finally, agreement between fTCD and VHF measures of laterality was investigated by administering both methods within the same participants (*Chapter seven*). Although moderate levels of concordance were found between dominance classifications obtained with these methods, correlations between their LIs were low and non-significant. This may be attributable to the VHF method's sensitivity to efficiency of interhemispheric transfer in addition to asymmetry of word processing.

Overall, the findings of this thesis demonstrate that a multidimensional view of lateralisation that considers the importance of interactions between the hemispheres is informative for understanding how processing of language may be altered in developmental disorders.

Table of Contents

Acknowledgements ..... *i*

Thesis Abstract ..... *iii*

Table of Contents ..... *vii*

Chapter 1 Introduction *1*

1.1 The meaning of language lateralisation **1**

1.2 Methods of measuring language lateralisation **1**

1.2.1 Clinical work 1

1.2.2 Behavioural measures 3

1.2.3 Brain imaging methods 5

1.3 Individual variability in lateralisation and language ability **8**

1.3.1 Is lateralisation beneficial? 9

1.3.1.1 Studies in typical language functioning 12

1.3.1.2 Studies in impaired language functioning 15

1.4 Interhemispheric interactions: the missing link? **20**

1.4.1 Mechanisms of intra- and inter-hemispheric interactions 21

1.4.2 Hemispheric interactions in language impairment 26

1.4.2.1 Differences in functional interhemispheric connectivity 26

1.4.2.2 Differences in structural interhemispheric connectivity 27

1.4.2.3 Behavioural evidence of weak interhemispheric connectivity 29

1.4.3 Interhemispheric connectivity and functional lateralisation 31

1.5 Variability in lateralisation of different language sub-processes **32**

1.5.1 Variability in lateralisation and language impairment 34

1.5.2 Variable lateralisation and weakened interhemispheric connectivity: the double-hit hypothesis 34

1.6 Summary and aims of the current thesis **35**

Chapter 2 Methodological considerations in assessment of language lateralisation with fMRI: a systematic review *41*

2.1 Introduction **42**

2.2 Materials and Methods **44**

2.2.1 Eligibility criteria 45

2.2.2 Search strategy and selection process 45

2.2.3 Data collection and data summary 46

2.3 Results **47**

2.3.1 Methods for calculating LI 47

2.3.2 Thresholding 49

2.3.2.1 Multiple thresholds 50

2.3.2.2 Variable thresholds 50
2.3.2.3 T-weighting and threshold independent methods 51
2.3.2.4 Bootstrapping 53
2.3.3 Activity measure 54
2.3.4 The flip method 57
2.3.5 Dominance classification 59
2.3.6 Effect of region of interest on laterality 64
2.3.6.1 Global and regional LIs 64
2.3.6.2 Regional approaches to LI calculation 66
2.3.6.3 Regional heterogeneity in laterality 67
2.4 Summary and conclusions 68
Chapter 3 Measuring language lateralisation with different language tasks: A systematic review 71
3.1 Introduction 72
3.2 Materials and Methods 75
3.2.1 Data analysis 75
3.3 Results 76
3.3.1 Verbal fluency 82
3.3.1.1 LI strength, reproducibility and robustness 82
3.3.1.2 Effect of region of interest 83
3.3.1.3 Effect of baseline task 84
3.3.2 Sentence generation 84
3.3.2.1 LI strength, reliability and robustness 85
3.3.2.2 Effect of ROI 85
3.3.2.3 Effect of baseline 86
3.3.3 Passive speech listening 87
3.3.3.1 LI strength, reliability and robustness 88
3.3.3.2 Effect of ROI 88
3.3.3.3 Effect of baseline task 88
3.3.4 Text reading 89
3.3.4.1 LI strength and effect of ROI 89
3.3.5 Phonemic judgement 90
3.3.5.1 LI strength, reliability and robustness 90
3.3.5.2 Effect of ROI 91
3.3.5.3 Effect of baseline task 91
3.3.6 Semantic decision 91
3.3.6.1 LI strength, reproducibility and robustness 92
3.3.6.2 Effect of ROI 93
3.3.6.3 Effect of baseline tasks 93
3.3.7 Sentence comprehension 94
3.3.7.1 LI strength, reliability and robustness 95
3.3.7.2 Effect of ROI 95

3.3.7.3 Effect of baseline task 96
3.3.8 Naming 96
3.3.8.1 LI strength, reliability and robustness 97
3.3.8.2 Effect of ROI 98
3.3.8.3 Effect of baseline task 98
3.3.9 Combined Task Analysis 99

3.4 Summary and conclusions 100

Chapter 4 Profile of language abilities in a sample of high functioning adults with developmental disorders 105

4.1 Introduction 106
4.1.1 Dyslexia in adulthood and higher education 106
4.1.2 Profile of abilities in adults with developmental disorders 108
4.1.3 Dyspraxia in adulthood 111
4.1.4 Summary of chapter aims 113

4.2 Methods 113
4.2.1 Participants 113
4.2.2 Procedure 115
4.2.3 Language assessment battery 116
4.2.3.1 Expression, Reception and Recall of Narrative Instrument (ERRNI) 118
4.2.3.2 Wechsler Adult Intelligence Scale (WAIS) IV: Digit span 118
4.2.3.3 Wechsler Abbreviated Scale of Intelligence (WASI): Vocabulary 119
4.2.3.4 York Adult Assessment Battery- Revised (YAA-R): Rapid naming 119
4.2.3.5 Test of Word Reading Efficiency (TOWRE) 120
4.2.3.6 NEPSY- A developmental neuropsychological assessment 120
4.2.3.7 Test of Adolescent and Adult Language: Sentence assembly 121
4.2.3.8 Communication Checklist self-report (CC-SR) 122
4.2.4 Normality of the measures 122

4.3 Aim 1: Characterisation of the profile of impairments in a sample of adults with developmental disorders 124
4.3.1 Statistical analyses 124
4.3.2 Results 127

4.4 Aim 2: Investigation of which measures were most sensitive for discriminating dyslexic from non-dyslexic participants 130
4.4.1 Statistical analyses 130
4.4.2 Results 132

4.5 Aim 3: Correspondence between diagnostic labels and data-driven sub-groups of participants 139
4.5.1 Statistical analyses 139
4.5.2 Results 141

4.6 General Discussion 143

4.6.1 Aim 1: Characterisation of the profile of impairments in a sample of adults with developmental disorders 144
4.6.2 Aim 2: Determination of which measures were most sensitive for discriminating dyslexic from non-dyslexic participants 145
4.6.3 Aim 3: Correspondence between diagnostic labels and data-driven sub-groups of participants 148
4.6.4 Limitations of the study 150
4.6.5 Summary and conclusions 153

Chapter 5 Inconsistency in lateralisation of language functions: a risk factor for language impairment? 155

5.1 Introduction 156
5.1.1 Defining 'atypical' laterality 156
5.1.2 Language lateralisation and language impairment 157
5.1.3 Aims of the current study 159

5.2 Methods 160
5.2.1 Materials and procedure 160
5.2.2 Measurement of language skills 160
5.2.3 Measurement of lateralisation 161
5.2.3.1 Phonological decision 163
5.2.3.2 Semantic decision 164
5.2.3.3 Sentence generation 164
5.2.3.4 FTCD analysis and calculation of laterality indices 165
5.2.4 Analysis plan 168
5.2.4.1 Pre-registered analyses 168
5.2.4.2 Non-preregistered exploratory analyses 170
5.2.5 Participants and statistical power analysis 172
5.2.5.1 Details of recruited sample 172
5.2.5.2 Summary of departures from pre-registered methods 173

5.3 Results 173
5.3.1 Behavioural data 173
5.3.2 LI values 174
5.3.3 Pre-registered Analyses 176
5.3.3.1 Consistency in lateralisation 176
5.3.3.2 Comparison of language measures between consistent and inconsistent groups 177
5.3.3.3 Consistency in lateralisation and handedness 177
5.3.4 Non-preregistered exploratory analyses 180
5.3.4.1 Correlations between task LIs 180
5.3.4.2 Consistency of lateralisation in the two groups 183
5.3.4.3 Consistency in lateralisation over time 183
5.3.4.4 Comparing strength of lateralisation 184

5.4 Discussion 186

5.4.1 Lateralisation and language impairment: Direction of causality 186
5.4.2 Patterns of laterality across language tasks 189
5.4.3 Inconsistency in lateralisation within a task 191
5.4.4 Further investigations using fMRI 191
5.4.5 Limitations of the study 193
5.4.6 Summary and conclusions 194

Chapter 6 Testing the interhemispheric deficit theory of dyslexia using the visual half-field technique 195

6.1 Introduction 196
6.1.1 Models of reading 196
6.1.2 The visual half-field paradigm 198
6.1.3 Developmental dyslexia and interhemispheric processing 201
6.1.4 Predictions of models 202
6.1.5 Aims of the current study: A replication 204

6.2 Method 205
6.2.1 Participants 205
6.2.2 Materials and Method 205
6.2.3 Hypotheses and analysis 208
6.2.3.1 Pre-registered analyses 208
6.2.3.2 Non-pre-registered analyses 209

6.3 Results 211
6.3.1 Preregistered analysis: Replication of comparison of RBA in dyslexics and controls 211
6.3.2 Non-pre-registered exploratory analyses 214

6.4 Discussion 216
6.4.1 Poorer LVF accuracy in dyslexia 216
6.4.2 Failure to replicate a reduced bilateral advantage in dyslexia 218
6.4.3 Methodological differences between the studies 220
6.4.4 Future directions 223
6.4.5 Summary and conclusions 224

Chapter 7 Correspondence between Doppler ultrasound and the visual half-field method as measures of language lateralisation 225

7.1 Introduction 226
7.1.1 Comparisons between different measures of laterality 227
7.1.2 The effect of inconsistent laterality on interhemispheric transfer 229
7.1.3 The effect of interhemispheric transfer efficiency on language abilities 230
7.1.4 Current study: Research Questions and predictions 232

7.2 Methods 233
7.2.1 Participants 233
7.2.2 Materials and method 233
7.2.2.1 Language assessment battery 233

7.2.2.2 Visual half-field task 233
7.2.2.3 fTCD tasks 234
7.2.3 Calculation of laterality indices and classification into dominance categories 235
7.2.4 Summary of measures 236
7.2.5 Statistical analyses 237
7.3 Results 239
7.3.1 Research question 1: Concordance and correlations between VHF and fTCD laterality measures 239
7.3.2 Research question 2: Effect of lateralisation consistency on VHF measures of interhemispheric transfer 242
7.3.3 Research question 3: Interhemispheric transfer efficiency as a protector against inconsistent laterality 244
7.4 Discussion 246
7.4.1 Summary of findings 246
7.4.2 Correspondence between VHF and fTCD measures of laterality 246
7.4.3 Effect of inconsistency in lateralisation on interhemispheric transfer 249
7.4.4 Preliminary test of the double-hit hypothesis 250
7.4.5 Limitations of the study 251
7.4.6 Summary and conclusions 252
Chapter 8 Overall Discussion 255
8.1 Language lateralisation and language impairment 255
8.2 Interhemispheric processing and dyslexia 257
8.3 The double-hit hypothesis of language impairment 258
8.4 Measurement of lateralisation with different methods 259
8.5 Concluding remarks 261
Bibliography 263
Appendix 1- Systematic literature search 289

Chapter 1 Introduction

1.1 The meaning of language lateralisation

Many lines of research have converged on the well-established finding that language processing in the brain shows a strong bias towards the left hemisphere, known as cerebral lateralisation for language. However, the phrase 'the brain shows a left lateralisation for language' has been used to mean many different things across different time points in research history. For something apparently so fundamental in the neurobiology of speech and language, 'lateralisation for language' remains a somewhat vague term used to refer to a collection of asymmetries across multiple levels measured in different ways. Asymmetries in behavioural performance, anatomical brain structure, functional brain activity, and brain connectivity are all seemingly interchangeably labelled as 'the lateralisation of language', without much progress as to the causal (or otherwise) relationships between these different levels of asymmetries. The meaning of language lateralisation at both a methodological and a theoretical level has been the subject of intensive study for over a century. The definition of language laterality appears to change according to the methodology used to measure it and the purpose of its measurement. This variability also reflects how conceptualisations of laterality have changed over time as different methods have become available.

1.2 Methods of measuring language lateralisation

1.2.1 *Clinical work*

The first reports of lateralisation for language emerged as a result of Marc Dax and Paul Broca's studies of patients with brain injury in the 19th century (Broca,

1861; Dax, 1865). They noticed that language impairments were strongly associated with damage to the left hemisphere, with right hemisphere damage typically leaving language intact. On the basis of this work, the left hemisphere was seen as critical for aspects of language processing for which the non-dominant right hemisphere is redundant. More recently, in preoperative mapping of the language system (e.g. prior to brain tumour surgery) lateralisation of language has been measured using the Wada test, in which temporary anaesthesia is induced in one hemisphere at a time by injection of sodium amobarbital in the corresponding internal carotid artery. This allows one to then probe the functions that depend on the anaesthetised hemisphere prior to neurosurgery, to avoid removal of cortex necessary for language. In this line of work, lateralisation is by necessity defined as a dominance of one hemisphere and a redundancy of the other for a particular function. However, some patients may still be classed as bilaterally organised for language, if injection of either hemisphere leads to (i) no deficit, (ii) partial deficit or (iii) complete loss of language (Bernal & Ardila, 2014). Indeed, there have been reports of patients who demonstrate dominance of different sides for different language functions on Wada testing (e.g. leftward expressive versus rightward receptive), although reports of such 'crossed dominance' cases are rare (Bernal & Ardila, 2014; Lee et al., 2008).

A new surge in interest and research on laterality of language was sparked by Sperry and Gazzaniga's seminal work with 'split-brain' patients in the 1960s. These patients underwent complete transection of the corpus callosum and anterior and hippocampal commissures (the white matter tracts connecting the two cerebral hemispheres) as treatment for intractable epilepsy. After such

disconnection, the isolated functioning of each hemisphere can be probed using lateralised stimulus presentation; since sensory fibres are typically crossed (most fully in the visual system), a stimulus presented to one side will be transmitted to the contralateral hemisphere only. Gazzaniga and Sperry (1967) found that such patients were unable to name pictures of objects presented in their left visual field and thus accessible only to the right hemisphere, typically responding that they did not see the object. Surprisingly however, such patients were able to reliably pick out the correct object from a selection using their left hand. This demonstrated that while the right hemisphere had perceived the object and extracted its shape and identity, it was unable to express this verbally; verbal output is instead reliant on the left hemisphere which truthfully reports that *it* saw nothing. Further extensive work with these patients provided a wealth of information on the language capabilities of the two hemispheres. Generally, this research suggested that while the right hemisphere could understand some semantic aspects of language, syntactic processing was constrained to the left hemisphere only (Gazzaniga & Hillyard, 1971). The right hemisphere was therefore described by Sperry (1982) as “largely mute and agraphic, but nevertheless able to comprehend, at a moderately high level, words spoken aloud”. In this way, this research made an important start to the process of mapping out the language faculties of the two hemispheres.

1.2.2 ***Behavioural measures***

The development of behavioural methods for measuring lateralisation made it possible to study lateralisation in non-clinical populations. Two dominant behavioural paradigms for measurement of language lateralisation that emerged around the 1960s and 70s were dichotic listening for the auditory domain and the

visual half-field (VHF) task for the visual domain. Both tasks rely on the crossing of fibres within these sensory systems (see Fig. 1.1 for schematic illustration). In the case of dichotic listening, different streams of auditory linguistic stimuli such as words are presented to each ear; participants are either asked simply to report as many words as they can or to attend to one ear selectively. In either case, a right ear advantage has been demonstrated in which participants report more verbal stimuli presented to the right ear than the left (Hugdahl, 2011; Kimura, 1961). Similarly in the VHF task, word stimuli are presented briefly to either the right or the left of fixation; words presented in the right visual field (which thus stimulate the left hemisphere) are typically recognised more quickly and accurately than words presented in the left visual field, termed the right visual field advantage (Hellige, 1993).

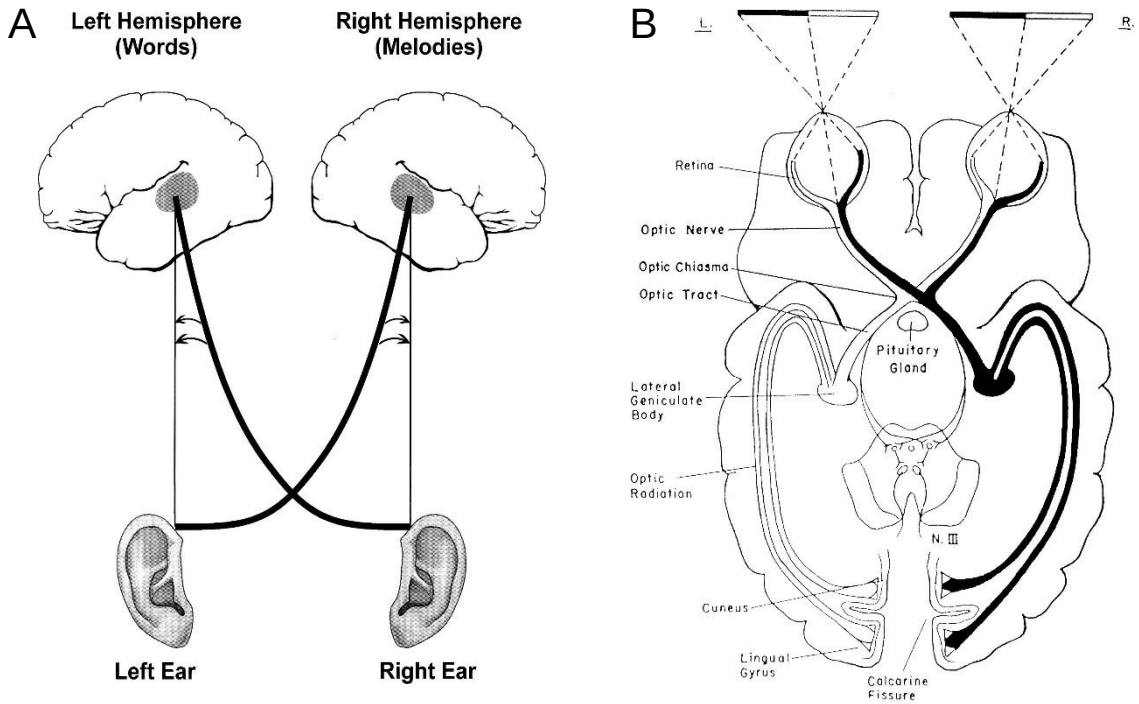


Figure 1.1: ***Dichotic listening and visual half-field paradigms.*** *(A) Schematic illustrating the partial crossing of auditory input presented to right and left ears. An advantage for reporting words is typically found for the left hemisphere (right ear input)*

whereas an advantage for melodies is typically found for the right hemisphere (left ear input). (B) Schematic illustrating the crossing of visual input presented to left and right visual fields, resulting in contralateral representation of each visual field in primary visual cortex. Reprinted from Brain and Cognition, vol 76 issue 2, D. Kimura, From Ear to Brain, pages 215-216, Copyright (2011), with permission from Elsevier.

Such asymmetries in behavioural performance are usually discussed with respect to laterality as a processing advantage in the dominant hemisphere; that is, processing of linguistic input/output in the left hemisphere is faster and more accurate or efficient than that in the right hemisphere. This allows for a more graded view of lateralisation in which there is a relative specialisation of the left hemisphere for language over the right. However, relating a performance measure with underlying functional specialisation of a hemisphere is not straightforward and multiple other factors could lead to lateralised differences in performance, such as attentional biases and cognitive strategies.

1.2.3 ***Brain imaging methods***

In the last few decades, laterality research has been transformed by the emergence of brain imaging methods, which allowed in vivo measurement of structural brain asymmetries. Post-mortem investigations had already led to the discovery of key asymmetries such as the right frontal and left occipital petalia; protrusions of the right frontal cortex and the left occipital cortex beyond their equivalents in the opposite hemisphere at each end of the brain. These relate to a general twisting of the two hemispheres known as the Yakovlevian torque that results in structures around the right Sylvian fissure being twisted forward relative to the left. The other early reported structural asymmetry is a leftward asymmetry of the planum temporale, first discovered by Geschwind and Levitsky (1968). The

planum temporale, which includes language-relevant regions including Wernicke's area, can be up to ten times larger on the left than the right (Steinmetz, 1996) with this asymmetry being present prenatally (Chi, Dooling, & Gilles, 1977).

Since their discovery, there has been much interest in the idea that such structural asymmetries may underlie the development of functional lateralisation for language. However, most studies have found no or weak evidence for a correlation between the planum temporale asymmetry and functional lateralisation (Dorsaint-Pierre et al., 2006; Tzourio-Mazoyer, Crivello, & Mazoyer, 2018). The advent of structural MRI techniques has allowed further structural asymmetries to be described, such as a leftward asymmetry of Broca's area in the inferior frontal gyrus and of Heschl's gyrus which houses the auditory cortex. More recently, a new structural asymmetry has been described by Leroy et al. (2015), termed the superior temporal asymmetrical pit or STAP. This refers to an asymmetry of the superior temporal sulcus which contains a region that is deeper on the right than the left. This depth asymmetry is present prenatally and throughout the lifespan, but was reported to be virtually absent in non-human primates. This study did not find, however, any relationship with functional lateralisation for language, being present in both left and right lateralised subjects in equal proportion. Overall, the significance of structural asymmetries to functional hemispheric specialisation for language remains unclear.

The development of brain imaging technologies furthermore allowed for more direct measurement of asymmetries in brain physiology in healthy human participants. Functional Magnetic Resonance Imaging (fMRI) uses measurements of oxygen levels in the blood supply to the brain as a proxy for

brain activity, and can be used to quantify the extent of activation across the two hemispheres. Methods for measuring laterality using this technique are reviewed in detail in Chapters two and three of this thesis. FMRI is by far the most dominant technique in laterality research, and has allowed great advances in our understanding of what aspects of language are lateralised. However the use of fMRI in some research scenarios is limited by the high cost of scans preventing large sample sizes, the unsuitability of the technique for use with participants with metal implants, pacemakers or claustrophobia, and the difficulties associated with scanning children.

An alternative method for measurement of laterality which overcomes these obstacles is functional transcranial Doppler ultrasonography (fTCD). This method uses ultrasound to measure the speed of blood flow in the middle cerebral artery (MCA) of each hemisphere. An ultrasound probe is attached on each side of the head at a region just in front of the ear where there is a thinning of bone in the skull, termed the temporal bone window. The phenomenon of Doppler shift allows calculation of the velocity of flow of red blood cells in this artery, due to the frequency change between the transmitted and reflected ultrasound signals. This can be used as a proxy for the neural activity in cortex supplied by this blood vessel due to tight neurovascular coupling; that is, an increase in blood flow velocity follows an increase in neural activity in order to maintain supply of resources to the neurons (Li & Freeman, 2015). The MCA territory covers regions within temporal, frontal and parietal cortices, including most key language areas (see Fig. 1.2). A laterality index can then be calculated based on the difference between the signal in the left and right channels (for a review, see Badcock & Groen, 2017). FTCD gives no information about localisation of activation within a

hemisphere, but offers a relatively inexpensive means of measuring laterality, and can be performed in samples for which fMRI measurement may be challenging.

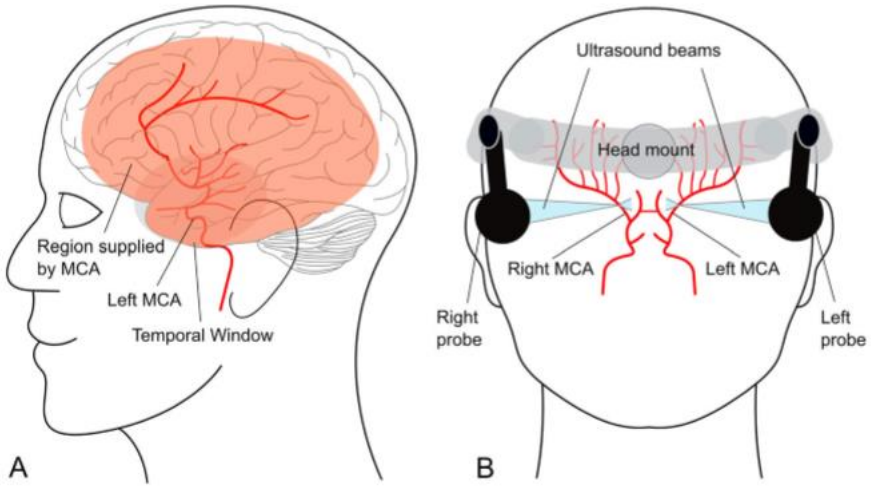


Figure 1.2: Territory of the middle cerebral artery. *(A) Schematic illustrates the range of cortical areas supplied by the middle cerebral artery. (B) Schematic illustrates the positioning of the fTCD probes and their insonation of the MCA. Figure reproduced with permission from Badcock and Groen, (2017).*

1.3 Individual variability in lateralisation and language ability

Neuroimaging work on lateralisation has led to a number of contemporary neurocognitive theories on the basis of lateralisation for language. These theories differ in the level at which they seek explanation for processing asymmetries, including lower level acoustic processing preferences (Poeppel, 2003), higher level linguistic aspects of language (Peelle, 2012) and the role of experience and learning in development (Yasuyo Minagawa-Kawai, Cristià, & Dupoux, 2011). Such theorising typically takes as its starting point the bias for left hemisphere organisation for language at a population level. That is, these theories typically do not consider the role or mechanisms of individual variability in lateralisation for

language. Evidence indicates that around 5-10% of the general population is right lateralised for language, with this figure being increased in left handed populations (Basic et al., 2004; Knecht et al., 2000; Mazoyer et al., 2014). Estimates of the incidence of bilaterality (equal engagement of both hemispheres during language) are complicated by concerns about the reliability of its measurement; however it has been reported to occur at similar levels to right lateralisation, and to again be increased in left- and mixed-handers (Basic et al., 2004; Kherdr, Hamed, Said, & Basahi, 2002; Mazoyer et al., 2014).

Lateralisation therefore presents a somewhat paradoxical phenomenon in which there is a strong bias towards left lateralisation at the population level, and yet a not insignificant level of variability across individuals. This contrasts to other cases of anatomical asymmetries in the human anatomy such as asymmetries in the organisation of the internal organs, which are rarely reversed; this so called 'situs inversus' disorder is only seen in 1 in 10,000 people (Torgersen, 1950). This raises the intriguing question as to the reason for this level of individual variability in language lateralisation, and whether it has any functional consequences.

1.3.1 *Is lateralisation beneficial?*

Theories on the origins of lateralisation typically make the explicit or implicit assumption that such asymmetric processing is beneficial for functioning. That is, it is assumed that the evolution of lateralised circuitry must have occurred because this offered some advantage in processing and behaviour that was then selected for. In an influential but controversial paper, Ringo, Doty, Demeter and Simard (1994) argued that lateralisation is the brain's solution to increasing brain size resulting in large interhemispheric transfer delays. This would render

processes that depended on integrated functioning of the two hemispheres too inefficient, creating a pressure for intrahemispheric circuitry to develop. Indeed, although there is evidence for changes in corpus callosum microstructure with increasing brain size that help to promote faster neural transmission, it is thought that these are not adequate to counteract the costs of longer interhemispheric transmission (Olivares, Montiel, & Aboitiz, 2001). Therefore it is likely that interhemispheric transfer in larger brained mammals did become less efficient, potentially forcing processes that required rapid interactive processing to be contained within a single hemisphere.

Alternatively, Rogers, Zucca and Vallortigara (2004) proposed that a division of labour across the hemispheres is beneficial as it enables simultaneous operation of different processes in the two hemispheres; this will result in an improved ability to dual-task. They provided a proof of principle of this proposal in a study of chicks. When exposed to light during incubation, chicks typically show different laterality patterns for different visual behaviours; specifically, searching for food is subserved by the left hemisphere, while checking for predators is conducted by the right hemisphere. Using a dual-task paradigm that simulated these two behaviours, lateralised chicks were found to perform both tasks better than chicks who were not lateralised (dark-incubated). This demonstrates how lateralisation can allow an organism to sustain simultaneous performance of two tasks that rely on different hemispheres. However, the relevance of such an animal model to human lateralisation can be questioned given that chick brain hemispheres are much less connected than the human cerebral hemispheres; hemispheric processing may therefore proceed in a more independent fashion similar to that of split-brain patients.

Others have suggested that lateralisation is simply a reflection of a packaging problem in the brain. Vallortigara (2006) argued that avoiding unnecessary duplication across the two hemispheres would allow an increase in neural capacity. Similarly, Corballis (2017) proposed that a pressure for lateralisation may have been created by the competing drives for a bigger brain but a restriction on the size of the head imposed by the feasible width of the birth canal in bipedal animals. Avoiding redundancy by having lateralised circuits would be beneficial for making the most of the space available to the brain within the constraints of skull size. However, Vallortigara (2006) pointed out that such accounts only explain what pressures could exist to promote absolute lateralisation, without offering insight into why strong biases towards a particular direction of such asymmetry would exist. The issue of why the left hemisphere specifically assumes a dominant role for language remains mysterious.

At a theoretical level, it seems reasonable to question why there would be such a strong bias towards left lateralisation for language if in fact it doesn't make any difference to language functioning. What model could account for such a strong bias whilst allowing a minority of cases where lateralisation is switched or absent to have no knock on effects on the functioning of the system? Some have argued that this pattern may be advantageous at a population level rather than an individual level. The potential advantages of coordinating direction of asymmetries across a population were discussed by Vallortigara (2006) in the context of asymmetries found in animals who live in groups; in shoals of fish for example, shared directions of behavioural asymmetries would be beneficial for the functioning of the community as a whole. In relation to human asymmetries such as handedness, Corballis (2017) conjectured that a population with

predominantly right handers but a minority of left handers may have constituted an optimal balance for behaviours such as warfare and sports.

The idea that lateralisation to the left hemisphere offers some benefit in processing at the individual level however has remained. This has led to much research investigating the functional consequences of atypical language laterality on language functioning. These can be broadly grouped into studies investigating this relationship in typically developing individuals, and studies investigating patterns of lateralisation in samples with impaired language; these will be reviewed in the following sections.

1.3.1.1 Studies in typical language functioning

Evidence in typically developing samples is mixed as to whether there are costs associated with having atypical language dominance. Early fTCD work by Knecht et al., (2001) did not find any detrimental consequences of atypical language lateralisation across a range of cognitive and linguistic abilities, including academic achievement, verbal fluency performance and proficiency in foreign languages. They concluded that atypical language lateralisation is not associated with any major behavioural consequences; that is, atypical lateralisation does not appear to confer a disadvantage in development. Early studies with fMRI similarly found no correlations between laterality indices and performance on standardised tests in samples of typically developing children (Holland et al., 2001; Wood et al., 2004).

More recent studies have however provided some evidence that weaker laterality is associated with poorer functioning. A positive relationship between LIs and language/reading performance was reported by Everts et al., (2009) using fMRI

and by Chiarello, Welcome, Halderman and Leonard (2009) using the VHF task. In a large sample of typically developing children, Groen, Whitehouse, Badcock and Bishop (2012) reported that those who were left lateralised for language (as assessed using fTCD) demonstrated better performance on measures of vocabulary and non-word reading than children who were not left lateralised. A study of adults by Mellet et al., (2014) reported that weak or bilateral organisation for language as measured by fMRI was associated with poorer performance across both verbal and non-verbal tasks compared to those showing strong lateralisation to either hemisphere. This suggests that it may not be the side of lateralisation that matters *per se*, but rather that language representation is focal to one hemisphere.

However, other studies have in fact reported a negative relationship between strength of laterality and language ability. In another VHF study, Hirnstein et al., (2010) reported an inverted U-shaped relationship between VHF LIs and mean accuracy on the same task, in which both extreme left and extreme right asymmetries were associated with poorer performance. Negative relationships have also been demonstrated between LIs and standardised measures of language ability using fMRI with a semantic paradigm (Bartha-Doering et al., 2018) and a sentence completion task (van Ettinger-Veenstra et al., 2010). This negative relationship may however be due to the nature of these tasks. In particular, the sentence completion task is known to recruit right hemisphere regions; thus, reduced lateralisation may not be indicative of atypical organisation for language in the canonical sense, but of adaptive recruitment of right hemisphere regions for sentence level processing. This task dependence of the direction of the laterality-language ability relationship is supported by a study by

Boles, Barth and Merrill (2008), who found relationships of different directions across a range of VHF and dichotic listening tasks. They proposed a neurodevelopmental model which attributed such task specific relationships to differences in the timing of maturation of lateralisation for different functions.

Other studies suggest that metrics of lateralisation other than strength may be relevant for predicting language ability. Kohler et al., (2015) studied lateralisation in a large sample of pre-school children using fTCD, and found that greater variability in the lateralised response of an individual (across trials within a single task) was associated with higher scores on standardised tests. They suggested that this increased variability may reflect a greater level of cortical plasticity that could be beneficial for development of language skills during the pre-school years. Furthermore, variability in laterality indices was found to be negatively correlated with age, suggesting stabilisation of lateralised responses as school age is approached. This highlights the possibility that the relationship between laterality and language ability may vary with developmental stage, rather than being consistent across the life-span. A further study by Gotts et al., (2013) suggests that differences in how the two hemispheres tend to communicate intra- and inter-hemispherically may be associated with performance outcomes. Using resting state fMRI, they demonstrated that left-lateralised language regions showed a preference for within hemisphere interactions, whereas right-lateralised ventral temporal regions showed a preference for between hemispheric interactions; critically, the more pronounced these preferences in an individual participant, the higher their scores on measures of vocabulary and visuospatial processing. This work demonstrates how research in this area should move

beyond simply looking at the degree of asymmetry of functional activation as a predictor of behavioural performance.

Overall, findings of studies on the laterality-ability relationship in typical language samples are mixed. One possible explanation for these inconsistencies are methodological differences across studies. In particular, studies vary in whether the ability measure used is performance on separately administered standardised assessments, or performance during the lateralising paradigm itself. In the case of behavioural tests, correlation of overall task performance (left plus right) with the extent of asymmetry (left minus right) represents a confound that is not always acknowledged by studies (see Hirnstein, Hugdahl, & Hausmann, 2014 for a discussion). Indeed, although at first glance the VHF studies by Chiarello et al., (2009) and Hirnstein et al., (2010) appear to report conflicting relationships, an appendix included by Chiarello et al., in fact reports the same negative relationship as Hirnstein et al., when correlating VHF asymmetry with mean naming performance on the same VHF task. Furthermore as discussed, different language processes may vary in the extent to which strong lateralisation is beneficial, and the extent to which increased right hemisphere recruitment should be interpreted as 'atypical organisation' for language.

1.3.1.2 Studies in impaired language functioning

The idea that atypical lateralisation for language may be associated with developmental disorders originated with the early work of Samuel Orton (1937). He proposed that weakly established cerebral lateralisation was a cause of disorders of spoken and written language, such as stuttering and dyslexia. He linked this to claims of a failure to establish strong right handedness and an increased persistence of 'mirror-reversal confusions' of letter writing and reading

in dyslexic children e.g. mistaking a ‘b’ for a ‘d’. Although contemporary evidence has generally failed to support either of these claims (although see Corballis, 2018), this theory has continued to influence research on the role of cerebral dominance in language ability. An influential theory of the genetics of handedness put forward by Annett (1985) further articulated these ideas. Her Right Shift Theory proposed that a lack of the so called ‘right-shift’ gene (RS+), which usually promotes left language lateralisation and right handedness, would be associated with increased risk of developmental dyslexia. In line with this theory, Smythe and Annett (2006) reported an association between reduced bias towards right handedness and poor phonological skills in a sample of primary school children. In general however, the claim that disorders of speech and language involve increased incidence of atypical handedness has received mixed empirical support, and effects are likely to be small if at all significant (Bishop, 1990; Eglinton & Annett, 1994). Furthermore, the relationship between handedness and lateralisation for language is not straightforward, with handedness typically being a poor predictor of lateralisation (Groen, Whitehouse, Badcock, & Bishop, 2013).

Since these early theories, the development of brain imaging technologies has enabled the link between atypical lateralisation and developmental disorders of speech and language to be measured more directly. An early fMRI study by Bernal and Altman (2003) reported that young children with speech delay demonstrated right sided language activation; however, a significant difference compared to a typically developing group was only found when looking at a sub-group of the sample over the age of three. The rationale for choosing to compare only children over this age (resulting in a very small sample size for the language delay group of only six children) is not clearly articulated in the paper. A more

recent fTCD study by Bishop, Holt, Whitehouse and Groen (2014) did replicate this finding however in a group of four year olds with language delay; children with current language difficulties demonstrated an absence of a bias towards left hemisphere language (with a mean LI not significantly different from zero). On the other hand, when the entire sample of impaired and unimpaired children was divided into those with typical and atypical lateralisation, no significant difference in language abilities was found. However, again this study is limited by a small sample size, with only 11 in the language problems group.

A number of studies have reported similar results for samples with a diagnosis of specific language impairment, now known as developmental language disorder (DLD). A small study of adults by Whitehouse and Bishop (2008) reported that a group with DLD demonstrated significantly lower laterality indices on fTCD than a range of other groups including those with a history of DLD (but no current language difficulties), a group with ASD and comorbid language difficulties and a group of typically developing adults. This reduced laterality has been further reported in fMRI studies of children with DLD using a range of language tasks (Badcock, Bishop, Hardiman, Barry, & Watkins, 2012; de Guibert et al., 2011). Strikingly however, a recent much larger scale study of over 100 twins with DLD and 150 typically developing twins reported no differences in lateralisation between groups using an fTCD animation description paradigm (Wilson & Bishop, 2018). This study represents a stark contrast to the rest of the literature, both in terms of its null result and its statistical power; studies reviewed so far have consisted of numbers of around 10 in each group. This strongly suggests that despite a seeming convergence of evidence, these previous significant findings are likely to represent false positives.

A similar story exists for the literature on asymmetries in dyslexia. Yamada et al., (2011) tracked the emergence of an adult-like reading network in children from the start of kindergarten to 3 months later. Children 'on-track' for normal literacy development were found to demonstrate a transition from a bilateral to a left lateralised reading network, involving a reduced recruitment of right hemisphere homologues. Conversely, a group of children at risk for reading difficulties showed an increase in bilateral recruitment in posterior temporo-parietal areas, along with additional recruitment of frontal areas. This suggests a failure to establish a left lateralised reading network. Xu, Yang, Siok and Tan (2015) similarly reported significantly reduced lateralisation in a group of Chinese speaking children with dyslexia compared to typically developing children. However, both these studies included rather small numbers of children ($n < 15$ in both dyslexic groups), and Xu et al., did acknowledge that the group difference did not survive correction for multiple comparisons. Significantly lower LIs for language have also been reported in adult samples with dyslexia, both on fMRI (Waldie, Haigh, Badzakova-Trajkov, Buckley, & Kirk, 2013) and fTCD (Illingworth & Bishop, 2009) measures.

Similar reports of reduced laterality have been found in other developmental disorders with relevance to speech and language. Although primarily a disorder of speech motor control rather than language, an absence of left lateralisation for language processing has been reported in both adults and children who stutter (Sato et al., 2011). A review of literature on structural and functional asymmetries in autism by Lindell and Hudry (2013) concluded that atypical asymmetries were common in autistic samples and linked to language impairments. A disruption to

cerebral lateralisation for language could potentially therefore represent a causal factor for language impairment that extends across diagnostic boundaries.

The publication of Wilson and Bishop's (2018) failure to replicate reduced laterality in a large sample of children with DLD brought to light the fragility of the seemingly convergent evidence from previous literature. The current review highlights some common practises across studies, such as the inclusion of very small samples, the use of uncorrected p values, and the potential for p-hacking offered by reporting comparisons of arbitrary sub-groups of samples. Aside from these methodological issues, a number of theoretical issues also remain. Studies differ in whether they report a significantly reduced level of lateralisation (i.e. a downward shift in the distribution of LIs such that lower levels of left lateralisation are common) or a significantly increased incidence of atypical dominance (with an absence of left dominant individuals), without much discussion of the theoretical distinction between these scenarios. Different possibilities as to the direction of causality between disrupted lateralisation and impaired language (see Fig. 1.3) are also often not acknowledged. As discussed by Bishop (2013), although interest has predominantly been in an ‘endophenotype’ model in which disruption to lateralisation mediates genetic risk for language impairment, alternative models include additive/interactive risks (genetic risk and weak laterality exert independent causal influences on language impairment), pleiotropy (weak laterality and language impairment share a common genetic cause but are not themselves causally related) and neuroplasticity (experience of development with impaired language itself causes development of weak laterality). Most research appears to assume an endophenotype model, with very

little discussion of possible alternative mechanisms that could underlie an observed relationship.

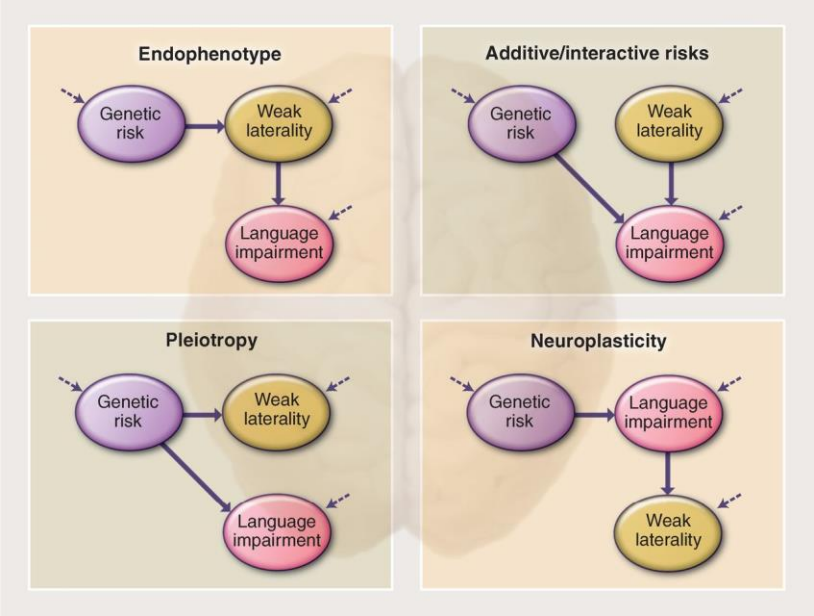


Figure 1.3: Different theoretical models of the relationship between genetic risk, weak laterality and language impairment. *Reproduced with permission from Bishop (2013).*

1.4 Interhemispheric interactions: the missing link?

In considering this evidence and how to progress in the study of this question, it is pivotal to have a clear concept of what it means to have 'atypical lateralisation'; that is, what are the necessary conditions required within a pattern of lateralisation for typical language functioning to develop? This will require moving beyond a simple typical-left versus atypical-right classification, to explore in more depth individual variability in profiles of lateralisation across different levels. That is, the same extent of functional asymmetry in brain activation measured in two individuals could be underwritten by distinct patterns of brain organisation. For example, recent research is considering how dynamic changes in intra- and interhemispheric interactions may form the mechanisms that lie behind

asymmetry of activation and function. This becomes particularly interesting in situations in which the brain may have to integrate outputs from operations that are lateralised to opposite hemispheres. This will require tightly regulated interactions between the two hemispheres in order to facilitate both their individual specialised processing and the final integration of their outputs. This is generally discussed in relation to cases where integration of lateralised functions from different domains is needed; however, it is also of interest to apply such discussions to cases where different sub-processes within the same domain lateralise to different hemispheres, increasing the need for frequent interhemispheric integration.

In the following sections I will discuss three key questions of interest that form the foundation for the investigations in my thesis. Firstly, how do the two hemispheres interact and how might this constitute a mechanism for hemispheric specialisation for a function? Secondly, how might such mechanisms of hemispheric interactions be disrupted in language impairment? And thirdly, how might these mechanisms be affected within a system in which different sub-processes are lateralised to different hemispheres?

1.4.1 *Mechanisms of intra- and inter-hemispheric interactions*

Two concepts are useful in considering how the two hemispheres may interact during lateralised activity. The first is the temporal coupling of processing within spatially separated areas, described as their functional connectivity. Of further importance is how activity in one area can have a causal influence on the activity of another area, known as their effective connectivity. There are many possible ways in which the brain may change its functional and effective connectivity both within and between the hemispheres during performance of a lateralised task.

Keller and Kell (2016) discuss two such possible models. Firstly, functional connectivity within the specialised hemisphere may be increased relative to the non-dominant hemisphere such that a within-hemisphere network is formed or strengthened (asymmetric intrahemispheric functional connectivity). Secondly, functional connectivity of an area located within the specialised hemisphere with areas in the non-dominant hemisphere may be increased relative to that of its homotopic area with areas in its respective contralateral hemisphere (asymmetric interhemispheric functional connectivity).

Stephan, Fink and Marshall (2007) further discuss the nature of inter-hemispheric interactions that may facilitate lateralised processing. The two hemispheres may cooperate in order for information transfer to occur (Liepmann, 1912; Poffenberger, 1912); for example, sensory information that enters first into the non-dominant hemisphere will have to be passed to the dominant hemisphere for further specialised processing. Alternatively, there may be competition between the hemispheres in the form of interhemispheric inhibition, with the dominant hemisphere inhibiting the processing and output of the non-dominant hemisphere (Chiarello & Maxfield, 1996; Kinsbourne, 1970). Lastly, it has been suggested that there may in fact be recruitment of the non-dominant hemisphere by the dominant hemisphere if the task is sufficiently complex and challenging (Belger & Banich, 1998).

These different possibilities in intra- and interhemispheric connectivity patterns during lateralised activity can be investigated using modern methods of modelling connectivity within the brain. These include dynamic causal modelling (DCM), structural equation modelling (SEM) and psychophysiological interactions analysis (PPI). The use of such methods in laterality research was reviewed by

Stephan et al., (2007) and has since grown. Seghier, Josse, Leff and Price (2011) used DCM on activity during a semantic decision task, and reported reduced effective connectivity from left to right dorsal frontal areas that was associated with increased left lateralisation; this relationship was specific to the use of word stimuli and not shown when picture stimuli were used for the same task. They proposed that this reflects reduced information transfer from left to right frontal language areas, which may be implemented by inhibition of the right hemisphere by the left to disrupt inter-hemispheric functional coupling. In counterpoint, a recent study by Keller and Kell (2016) using PPI reported increased functional connectivity between right occipital cortex and left lateralised language regions during a reading task, which they interpreted as increased transfer of sensory information to the specialised hemisphere for further processing. Keller and Kell (2016) further reported greater functional connectivity between regions in the left language network than that shown between right homotopic regions. This work suggests that lateralized processing is associated with increased transfer of sensory information from the non-dominant to the specialised hemisphere (asymmetric interhemispheric connectivity), and increased functional connectivity within the specialised hemisphere relative to that in the non-dominant hemisphere (asymmetric intrahemispheric connectivity).

Indeed, it has been suggested that the two hemispheres may possess inherent biases for different patterns of connectivity that could form the basis of left lateralisation of function. A large scale meta-analysis of neuroimaging studies of the language network by Vigneau et al., (2011) reported that while the majority of left hemisphere activity peaks were unilateral (with no homotopic activation present), the majority of right hemisphere peaks were bilateral. Resting-state

studies have further reported higher functional connectivity between left hemisphere language areas than between their right hemisphere homologues, which instead showed greater functional connectivity with left default network areas (Gotts et al., 2013; Raemaekers, Schellekens, Petridou, & Ramsey, 2018). This has led to suggestions that the left hemisphere may have a preference for intrahemispheric interactions, whereas the right may be biased towards more interhemispheric interactions. Further, a number of studies report that the greater these asymmetries in intrahemispheric resting connectivity, the greater the degree of functional language lateralisation shown (Joliot, Tzourio-Mazoyer, & Mazoyer, 2016; Liu, Stufflebeam, Sepulcre, Hedden, & Buckner, 2009; Raemaekers et al., 2018; Wang, Buckner, & Liu, 2014). These biases in intrinsic connectivity could either be a consequence of functional asymmetries or indeed be instrumental in the establishment of lateralisation of function.

Biases in connectivity patterns between the hemispheres have also been reported in measures of interhemispheric transfer times (IHTT). Estimates of the time taken for a visual stimulus to be transferred from the initially stimulated contralateral hemisphere to the ipsilateral hemisphere can be obtained using ERP measurements of IHTT (Nowicka & Tacikowski, 2011). Here, after lateralised visual stimulus presentation, IHTT is taken as the delay between the latencies of a particular ERP component measured from the contralateral (earlier response) and ipsilateral (later response) hemisphere is measured.. Using this method with a range of visual stimuli, it has been reported that an asymmetry in IHTTs is apparent in the form of significantly shorter IHTTs from the right to the left hemisphere than vice versa (Brown, Larson, & Jeeves, 1994). Complementary MEG evidence from Barca et al., (2011) reported that after uni-

lateral presentation of word stimuli, greater spread of activation from right to left visual cortex was seen after left visual field stimulation than that seen from left to right visual cortex after right visual field stimulation. Interestingly, this pattern of asymmetry in IHTTs has only been reported in right handers, with Iwabuchi and Kirk (2009) finding no such asymmetry in left handers (balanced IHTTs in both directions). This has led many to suggest that such a pattern may have relevance to hemispheric lateralisation, given the increased incidence of atypical lateralisation in left handers.

Some have argued that this typical pattern of asymmetry in IHTT is underpinned by structural differences in the corpus callosum (Marzi, Bisiacchi, & Nicoletti, 1991), with reports of an increased number of fibres in the splenium traversing from the right hemisphere to the left than in the reverse direction (Putnam, Wig, Grafton, Kelley, & Gazzaniga, 2008). However, this anatomical interpretation of asymmetries in IHTT cannot account for reports that this asymmetry depends on the nature of the stimulus. An early study by Nowicka, Grabowska and Fersten (1996) compared IHTT asymmetries for verbal versus non-verbal visual stimuli; word stimuli showed the typical shorter latencies for IHTTs from the right to the left hemisphere, whereas for grating stimuli latencies were shorter from the left to the right. More recently using DCM analysis, Stephan, Penny, Marshall, Fink and Friston (2005) reported that during performance of two decision tasks with left visual field stimuli, an enhancement of effective connectivity from the right lingual gyrus to its left homotope was only seen for a linguistic task, and not when a non-linguistic task was performed on the same stimuli. This research suggests that interhemispheric transfer could be seen not as a slave to underlying static anatomical differences, but as a dynamic process in which the brain can

adaptively modulate patterns of interhemispheric transfer according to task demands.

1.4.2 *Hemispheric interactions in language impairment*

As reviewed previously, evidence for differences in the strength of functional language asymmetries in language impaired samples is mixed and inconclusive. Interhemispheric communication and connectivity may provide an additional dimension of analysis to tease apart differences between patterns of lateralisation seen in language impaired and non-impaired individuals. In order for this research to develop further, it will be important to consider how language impairment may be associated with (i) differences in lateralisation across different language sub-processes and (ii) disruption to mechanisms of intra- and inter-hemispheric interactions.

1.4.2.1 Differences in functional interhemispheric connectivity

There is indeed a body of work demonstrating differences in functional and structural connectivity both within and across the hemispheres in language impairment. Language and reading require the integration of processing across multiple distributed regions, and it has been proposed that language impairment may be associated with weaker connectivity between such areas (Paulesu et al., 1996). This has been particularly explored in the case of dyslexia. An early study by Horwitz, Rumsey, and Donohue (1998) reported that while typical readers showed evidence of leftward asymmetric intrahemispheric connectivity, this was absent in dyslexics; specifically, greater functional connectivity was seen between the left angular gyrus and left visual and language areas than between their right homologues in typical readers only. Similarly and more recently, van der Mark et al. (2011) reported reduced functional connectivity between the left

ventral occipitotemporal cortex (also known as the visual word form area) and left language areas (inferior frontal gyrus and inferior parietal lobule) in a group of dyslexic children compared to control children. This suggests a failure to establish a strong left lateralised linguistic processing network in dyslexic readers. Conversely, Finn et al. (2014) reported increased functional connectivity between right hemisphere language homologues in children with dyslexia compared to typically developing children. Visual areas were more functionally coupled with the right IFG and angular gyrus, and the right supramarginal gyrus demonstrated increased connectivity with both ipsilateral and contralateral areas. This evidence suggests that the normal left biased asymmetric intrahemispheric and interhemispheric connectivity patterns are diminished in dyslexia.

1.4.2.2 Differences in structural interhemispheric connectivity

Differences in the structure of white matter pathways have also been reported in language impairment. In a review of diffusion-tensor imaging (DTI) studies of dyslexia, Vandermosten, Boets, Wouters and Ghesquière (2012) concluded that the most consistent finding across studies was a reduced fractional anisotropy (FA) in a left temporoparietal region in dyslexics relative to controls, indicative of weaker white matter microstructure (e.g. reduced myelination). This has generally been interpreted as disruption to the arcuate fasciculus, a white matter tract that runs underneath this region. An interesting alternative interpretation however has been suggested in a recent study by Zhao, Thiebaut de Schotten, Altarelli, Dubois, and Ramus (2016). Using spherical deconvolution tractography to specifically identify the arcuate fasciculus, they indeed found evidence of reduced FA in the left arcuate of children with dyslexia related to typically developing children; however, this was not found when an alternative measure of

white matter tract density was used called hindrance-modulated oriented anisotropy (HMOA). This differs from FA primarily in its lack of susceptibility to the effects of crossing fibres which can spuriously reduce FA values. They argued that this group difference in FA of the arcuate fasciculus may in fact reflect increased crossing of fibres in this region that could represent an increase in interhemispheric connections in the corpus callosum.

Indeed, abnormalities of the structure of the corpus callosum have been reported in language impaired samples; however this evidence is not always consistent. Some studies report a negative relationship between corpus callosum FA and reading ability, with an increase in corpus callosum FA being associated with poorer reading (Dougherty et al., 2007; Frye et al., 2008). Such findings in the dyslexia literature have been interpreted as evidence for increased interhemispheric communication due to reduced hemispheric specialisation (Vandermosten et al., 2012). Conversely, other studies have reported a decrease in corpus callosum FA in language impaired samples (Hynd, Hall, Novey, & et al, 1995; Kim et al., 2006; Welcome & Joanisse, 2014). Indeed, in the study by Frye et al (2008), while orthographic skills were negatively associated with corpus callosum FA, phonological awareness showed a positive relationship in the typical readers group (but not in the dyslexic group).

More careful interpretation of the evidence reported by Dougherty et al. (2007) in fact reveals compatible findings. They reported a negative relationship between radial diffusivity in the temporal segment of the corpus callosum and phonological awareness, which was interpreted as evidence that good readers showed reduced interhemispheric connectivity. Radial diffusivity refers to diffusivity in the direction perpendicular to the longitudinal axis of a tract, and is reduced by the

presence of myelin sheaths and axonal membranes. Thus, one plausible interpretation of an increased radial diffusivity would be the presence of a smaller number of large diameter axons. Whilst Dougherty et al (2007) appear to assume such a pattern would reflect reduced interhemispheric connectivity, in fact it has been shown that larger diameter axons allow faster transmission of neural impulses, and so their presence in the corpus callosum would be critical for efficient and fast interhemispheric communication (Aboitiz, López, & Montiel, 2003; Horowitz et al., 2015). Therefore, their finding of increased radial diffusivity associated with better reading related skills is consistent with more efficient interhemispheric transfer in good readers.

This idea is corroborated by evidence from a study by Carreiras et al., (2009) in which the acquisition of literacy in adulthood was associated with increased white matter in the splenium of the corpus callosum relative to adult illiterates. Acquisition of reading skill in a typically developing system (regardless of when that skill acquisition takes place) may thus be associated with strengthening of interhemispheric tracts in the corpus callosum that fails to occur in a system that yields impaired reading. The relevance of interhemispheric connectivity to language functioning is further suggested by reports of reduced connectivity within the anterior commissure and corpus callosum in adolescents born prematurely with language impairments, compared to pre-term adolescents with typical language abilities and full term controls (Northam et al., 2012).

1.4.2.3 Behavioural evidence of weak interhemispheric connectivity

This idea of weaker interhemispheric transfer in language impairments is also supported by behavioural evidence from visual half-field and dichotic listening paradigms. Damage to the corpus callosum has been reported to result in even

poorer performance for left-presented stimuli and thus a larger right visual field/ear advantage (Mohr, Pulvermüller, Rayman, & Zaidel, 1994; Musiek & Weihing, 2011). It is of interest then that studies of language impaired individuals have reported alterations of the normal patterns of right versus left performance in these paradigms. Using dichotic listening, Hakvoort et al., (2016) reported that dyslexic children with familial risk for dyslexia showed a larger right ear advantage than controls and non-dyslexic children with familial risk for dyslexia, caused by reduced left ear reports. Similarly in the visual half-field literature, a greater right visual field advantage has been reported in dyslexic samples. Henderson, Barca, and Ellis (2007) found poorer left visual field performance in a sample of dyslexic adults, and further an absence of the 'redundant bilateral advantage', a phenomenon in which simultaneous presentation of the same word to both visual fields results in improved recognition beyond that of right visual field words. The absence of this effect in split brain patients (Mohr et al., 1994) suggests that this can be interpreted as evidence of weakened callosally-mediated interhemispheric interactions.

ERP measurement of IHTTs in dyslexic samples suggests that this altered interhemispheric communication may relate to disruption of normal patterns of asymmetry rather than a general deficit in interhemispheric transfer. Martin, Thierry, Démonet, Roberts and Nazir (2007) reported that dyslexic participants failed to show the normal asymmetry in IHTTs for written word stimuli, instead showing balanced latencies for transfer of this visual verbal information in both directions. Thus, the ability of the brain to modulate the efficiency of interhemispheric transfer according to task demands may be affected in dyslexia.

The overall picture of research thus suggests that normal mechanisms of intra- and interhemispheric interactions may be disrupted in some way in individuals with language impairments. However, it should be noted that the apparent convergence in evidence for white matter alterations in dyslexia did not hold in a recent meta-analysis by Moreau, Stonyer, McKay and Waldie (2018). Aggregating data across studies, they failed to find any clusters that showed significant differences in dyslexics or significant correlations with reading ability. They attributed this to the failure of many studies to adequately correct for multiple comparisons when using DTI and the high level of inconsistency in the coordinates reported by different studies. Others have argued that the voxel based morphometry analysis approach ignores the possibility of heterogeneity within language impaired samples, and that instead of looking for isolated areas of overlap one should consider weakness in the network as a whole (Bathelt, Gathercole, Butterfield, & Astle, 2018). As they stand, DTI studies may therefore not currently constitute solid evidence for structural differences in interhemispheric tracts in language impairment; however behavioural evidence from VHF and dichotic listening paradigms and IHTT evidence from ERP paradigms still point to potentially interesting differences in interhemispheric processing.

1.4.3 *Interhemispheric connectivity and functional lateralisation*

How might differences in interhemispheric communication impact on functional lateralisation for language? Interestingly, a number of studies have reported a positive relationship between measures of corpus callosum macro- and micro-structure (e.g. surface area, volume, FA) and functional lateralisation for language (Josse, Seghier, Kherif, & Price, 2008; Putnam et al., 2008;

Westerhausen, Kreuder, Santos Sequeira, et al., 2006). This greater interhemispheric connectivity may give rise to greater functional lateralisation by facilitating better inhibitory control of the non-dominant hemisphere by the dominant hemisphere; in contrast, when interhemispheric connectivity is weak, the non-dominant hemisphere can proceed unregulated, resulting in more (and potentially maladaptive) bilateral activity. Conversely, an EEG study by Friedrich et al. (2017) reported a negative correlation between interhemispheric latency differences in the N1 component during dichotic listening and FA values in the posterior corpus callosum. This negative relationship was also reported in a recent meta-analytic paper by Karolis, Corbetta and Thiebaut de Schotten (2019) who demonstrated that regions with stronger functional lateralisation demonstrated reduced micro- and macrostructural connectivity within the corpus callosum. This evidence is more consistent with the idea that reduced interhemispheric communication results in stronger lateralisation of processing since the two hemispheres are more isolated from one another and thus activation does not spread as readily across them. This discussion of inconsistent evidence on the relationship between the structure of the corpus callosum and functional language lateralisation highlights the insufficiency of our understanding of the role of this structure in mediating interhemispheric interactions.

1.5 Variability in lateralisation of different language sub-processes

Performance of a single task will likely draw on specialised processing lateralised to different hemispheres. This will thus require efficient integration of such distributed processing across the hemispheres. Generally, this is considered in relation to laterality differences between different domains, such as differences

between lateralisation of language (typically left) and visuospatial processing (typically right). However, there is also evidence that different sub-components within a domain may show variability in their lateralisation. In the domain of language, aspects of language structure such as phonology are generally ascribed to the left hemisphere (Hickok & Poeppel, 2007), while more pragmatic aspects of language such as processing of prosody and metaphor are attributed to the right hemisphere (Bryan, 1988). However, there is further evidence that typically left lateralised sub-processes within language can lateralise to different hemispheres in certain individuals. For example, cases have been reported in which individuals show one pattern of asymmetry within frontal areas for expressive language but the reverse pattern of asymmetry within more posterior areas for receptive language (Häberling, Steinemann, & Corballis, 2016; van der Haegen, Cai, & Brysbaert, 2012).

Given such cases of intra-individual variability in lateralisation, Tailby, Abbott and Jackson (2017) have argued that attempts to determine “which is the dominant hemisphere?” for a given individual incorrectly assume homogeneity in laterality within the language network. This highlights a move in recent research towards a more multi-dimensional view of language lateralisation, in which different sub-components of the language system may lateralise independently (Bernal & Ardila, 2014; Bradshaw, Bishop, & Woodhead, 2017). Cases of fully crossed lateralisation are rare in the reported literature, possibly due to studies only reporting group level results or using a single lateralising task or region of interest; however, their existence raises questions as to the potential effects on interhemispheric interactions. In such individuals, even just a moderately complex linguistic operation will likely depend on sub-processes that are

lateralised to opposite hemispheres, thus placing increased demands on mechanisms of interhemispheric integration.

1.5.1 *Variability in lateralisation and language impairment*

Bishop et al. (2014) have proposed that such increased variability in the side of lateralisation of different language sub-processes could constitute an aetiological risk factor for development of language impairment. In their 'Left Brain Bias model', two phenotypes are proposed of either a left brain bias or no bias to left hemisphere language. In the former, development of language laterality is biased towards the left hemisphere, for example with a weighting of 90:10; conversely, in the no bias phenotype, language laterality has equal likelihood of developing left or right bias. In this model, development of lateralisation for each aspect of language functioning can be considered as an independent probabilistic process that either occurs with this systematic left bias, or without any such bias. Consequently, this model predicts that individuals with the no-bias phenotype will show more variability in the side of lateralisation of different language functions.

1.5.2 *Variable lateralisation and weakened interhemispheric connectivity: the double-hit hypothesis*

Having different language processes located in different hemispheres will necessarily increase demands on interhemispheric integration via the corpus callosum. Thus, in such a distributed network, there will be increased need for interhemispheric cooperation. In a normally developing brain this may not necessarily have any detrimental consequences for language function, as normal mechanisms of interhemispheric interactions can support good language functioning. However, given evidence of disruption to the corpus callosum in language impairment, such mechanisms of interhemispheric interactions may be

disrupted in individuals with language disorders. I hypothesise that language impairment may involve a 'double-hit' of (i) increased variability in the side of lateralisation of different language sub-processes (i.e. the no-bias phenotype) and (ii) weaker mechanisms of interhemispheric interactions as evidenced by alterations to the macro- and micro-structure of the corpus callosum. In this way, having language functions located in different hemispheres increases demands for interhemispheric integration on an already weakened system, resulting in impaired language functioning.

An alternative view could conceptualise increased distribution of language functions across the hemispheres as a consequence of weakened inter-hemispheric interactions. Weakened callosal connectivity may disrupt interhemispheric inhibition, resulting in dysregulation of simultaneous and potentially conflicting outputs from both hemispheres. This would lead to increased inconsistency as to which hemisphere is dominant on a given measure of language lateralisation. This failure of suppression of the non-dominant hemisphere by the dominant hemisphere due to weakened cross-callosal tracts could underlie the increased functional connectivity observed by Finn et al. (2014) between right language areas in dyslexic readers.

1.6 Summary and aims of the current thesis

This thesis aims to advance our understanding of the potential disruption to lateralisation in language impairment by considering the importance of consistency in lateralisation across language functions within individuals. It seeks to integrate these two literatures on the relationship between atypical laterality and language ability on the one hand, and the role of interhemispheric interactions in lateralisation and language functioning on the other. In so doing, it

incorporates a modern perspective on the mechanisms underlying lateralisation into research on the long standing question of the involvement of lateralisation in language disorders.

This thesis begins by considering how one can optimally measure lateralisation for language. In the next two chapters, I provide a systematic review of fMRI studies of language lateralisation from the last two decades, with a focus on methodological aspects of laterality measurement. This was motivated by the high level of variability in the methods used across studies, which can make comparisons across the literature difficult. The first systematic review (*Chapter 2*) considers the different protocols that have been used to quantify and classify laterality from fMRI data by studies published between 2000 and 2016. In reviewing this literature, I aimed to:

1. Consider whether the field is converging on common methods for assessing language lateralisation.
2. Evaluate the robustness and reliability of these different methods in order to make recommendations that could be used to guide methodological decisions in future laterality research.

The second systematic review (*Chapter 3*) provides a complementary review of the same literature, but with a focus on the language activation paradigms used for laterality measurement. Differences in strength of laterality have been reported across different paradigms, however it is not always clear whether these are attributable to true differences in lateralisation for different language processes, or merely to trivial methodological differences. In reviewing the different language paradigms used in the laterality literature, I aimed to:

1. Evaluate the strength, reliability and robustness of laterality measurements obtained with different language tasks.
2. Consider the effect of different methodological parameters on laterality measurements, including the choice of baseline task, region of interest and modality of the stimuli.
3. Make recommendations for optimal task design that can be used to guide future fMRI studies of language lateralisation.

In my first experimental study, I recruited a large group of adults (n = 104) with varying language abilities; the majority of whom had a diagnosis of a developmental disorder relevant to language (n = 67), and the remainder of whom had no such diagnosis (n = 37). This group was tested with a comprehensive language assessment battery, and with fTCD for measurement of lateralisation across three language tasks. In Chapter four, I present an analysis of the language assessment data from this adult sample. I aimed to:

1. Provide a characterisation of the profile of impairments demonstrated by this sample of adults with developmental disorders in order to establish that this sample demonstrated impaired language abilities.
2. Determine which measures were most sensitive for discriminating individuals with a diagnosis of dyslexia from those without.
3. Explore the data-driven similarity structure within this language assessment data in order to consider whether data-driven sub-groups correspond well to diagnostic categories.

In my second experimental study (*Chapter 5*), I analysed the fTCD laterality measurements taken with this same sample to consider the relationship between

within-individual variability in lateralisation for different language functions and language abilities. All 104 participants were tested with three fTCD language paradigms that aimed to engage different aspects of language processing; namely, phonology, semantics and speech production. Rather than taking a traditional approach in which the strength of lateralisation for a single language task is compared between an impaired and non-impaired group, this study took a different approach by assessing language skills in a continuous fashion and considering the importance of consistency in lateralisation across different language functions. As discussed previously, increased variability in the side of lateralisation of different language functions may result in a less efficient language network and reflect disruption to normal mechanisms of interhemispheric interactions. This study aimed to:

1. Test the hypothesis that an inconsistent pattern of lateralisation across the three language tasks would be associated with poorer language skills compared to a consistent lateralisation pattern.
2. Compare the incidence of consistent and inconsistent laterality patterns between developmental disorder and control groups.
3. Consider the relationship between consistency in lateralisation and handedness.
4. Compare consistency in lateralisation over time within performance of a single task between developmental disorder and control groups.

My third study (*Chapter 6*) tested a group of 47 dyslexic adults and 43 typical readers using a visual half-field paradigm in order to examine interhemispheric processing. Although more direct measures of corpus callosum microstructure such as those obtained via DTI would have been preferable to test for disrupted

interhemispheric connectivity in language impaired samples, the visual half-field technique offers a more easily administered behavioural method for assessing the functional dynamics of interhemispheric transfer of visual information during word recognition. Previous work has reported that dyslexic individuals fail to show the expected bilateral advantage in word recognition (L. Henderson et al., 2007), indicative of disrupted interhemispheric communication. My study aimed to compare visual half-field performance between dyslexic participants and controls in order to:

1. Replicate previous findings of an absence of this bilateral advantage in dyslexic participants.
2. Test for a difference in the magnitude of behavioural asymmetries in word recognition between dyslexic and control participants.
3. Test for a relationship between visual half-field measures of asymmetry and reading ability.

My final study (*Chapter 7*) combined measures of laterality obtained from fTCD and VHF methods within the same individuals (n = 58). Understanding how asymmetries measured from imaging and behavioural techniques relate to each other has both practical and theoretical implications. While some have argued that the VHF method represents a reliable alternative to fMRI for prediction of cerebral lateralisation for language (Hunter & Brysbaert, 2008), other studies have reported surprising low levels of correspondence between imaging and VHF laterality measures (Krach, Chen, & Hartje, 2006). This may reflect the VHF method's additional sensitivity to the efficiency of interhemispheric transfer of visual word information, on top of cerebral lateralisation for language. Further, by combining my VHF and fTCD laterality data, I was able to conduct a preliminary

test of my double-hit hypothesis that inconsistency in the side of lateralisation of different language functions *and* weakened interhemispheric connectivity together are the necessary conditions for increased risk of language impairment. Again, although this would have been optimally tested by collecting DTI measures in the same individuals as functional laterality measures, this was outside the scope and resources of my DPhil. Instead, the availability of fTCD and VHF data within the same individuals allowed me to conduct a preliminary test of this hypothesis. By combining fTCD and VHF laterality measures, I aimed to:

1. Determine whether fTCD and VHF methods yield dominance classifications within the same individuals that demonstrate a good level of agreement, and laterality indices that are significantly correlated.
2. Consider whether consistently and inconsistently lateralised individuals (as measured by fTCD) demonstrate significant differences on VHF measures of interhemispheric transfer.
3. Test a prediction of the double-hit hypothesis that strength of interhemispheric connectivity (as indirectly indicated by VHF measures) would have a significant effect on language abilities only in individuals who demonstrate inconsistent lateralisation across language functions, and not in individuals consistently lateralised for language.

My thesis ends with an overall discussion of the findings of my thesis (*Chapter 8*) and their implications for how we measure lateralisation, how we understand the relationships between different measures of asymmetries, and how our conception of the very nature of cerebral lateralisation may need updating to a multidimensional framework.

Chapter 2 Methodological considerations in assessment of language lateralisation with fMRI: a systematic review

Abstract

The involvement of the right and left hemispheres in mediating language functions has been measured in a variety of ways over the centuries since the relative dominance of the left hemisphere was first known. Functional magnetic resonance imaging (fMRI) presents a useful non-invasive method of assessing lateralisation that is being increasingly used in clinical practice and research. However, the methods used in the fMRI laterality literature currently are highly variable, making systematic comparisons across studies difficult. This chapter provides a systematic review of the different methods of quantifying and classifying laterality that have been used in fMRI studies since 2000, with the aim of determining which give the most robust and reliable measurement. Recommendations are made with a view to informing future research to increase standardisation in fMRI laterality protocols. In particular, the findings reinforce the importance of threshold-independent methods for calculating laterality indices, and the benefits of assessing heterogeneity of language laterality across multiple regions of interest and tasks. This systematic review was registered as a protocol on Open Science Framework (https://osf.io/hyvc4/) and has been published as a paper (Bradshaw, Bishop, et al., 2017).

2.1 Introduction

Non-invasive techniques for assessment of language lateralisation make it possible to probe the characteristics of language lateralisation in neurologically intact populations. Functional magnetic resonance imaging (fMRI) is a prominent non-invasive method that has been used to assess laterality. A laterality index (LI) is calculated based on a comparison of activation measures from each hemisphere, according to the following formula:

$$LI = \frac{L - R}{L + R}$$

This calculates laterality as the difference between activity in each hemisphere (L and R) divided by the total activity across the hemispheres. The LI gives a single value indicating the relative strength of left and right hemisphere activation for an individual. LI measurement may be required for clinical purposes in order to establish an individual's hemispheric dominance for language prior to surgery, as in patients with intractable epilepsy. Alternatively, a study may measure an LI to assess the strength or variability in lateralisation for a given language function in order to make inferences about the neural organisation of the language system. That is, studies may vary in whether the aim of LI measurement is to classify or to quantify lateralisation. This will have important implications for which methods of LI calculation are optimal for laterality measurement.

Interpretation of fMRI lateralisation research has been problematic due to a lack of standardisation of fMRI laterality protocols. Multiple arbitrary decisions must be made when calculating the L and R terms for use in the LI equation which might affect the LI value obtained (Jansen et al., 2006; Seghier, 2008). Such

variability in methodology can preclude systematic study of language lateralisation.

For example, when calculating an LI from active voxels in each hemisphere or region of interest (ROI), a decision must be made as to the threshold *p* value at which to view and analyse the images. Multiple studies have documented the dependence of the LI obtained on the threshold chosen (Abbott, Waites, Lillywhite, & Jackson, 2010; Adcock, Wise, Oxbury, Oxbury, & Matthews, 2003; Nadkarni et al., 2015; Rutten, Ramsey, van Rijen, & van Veelen, 2002; Seghier et al., 2004). As illustrated in Figure 2.1, as the threshold value is increased, the number of voxels surviving thresholding decreases, typically leading to an increase in the LI. Ultimately, above a certain threshold, no active voxels will remain in the non-dominant hemisphere, resulting in an LI of 1; and below a certain threshold many voxels will survive across both hemispheres, resulting in an LI of 0. Indeed, there are even reports of individuals whose LI shows a switch in dominance with a change in threshold level (Jansen et al., 2006; Ruff et al., 2008; Suarez, Whalen, O'Shea, & Golby, 2008; Wilke & Lidzba, 2007).

This illustrates just one preliminary issue that must be addressed when considering how to quantify lateralisation from fMRI data. Further decisions have to be made as to tasks used in an activation paradigm, whether the analysis focuses on a specific region of interest (ROI) or the whole hemisphere, and whether the quantification of activation is based on magnitude or extent of activity. If the LI is used to categorise individuals as left-, bilateral or right-lateralised, a suitable cut-off for categorisation must also be determined.

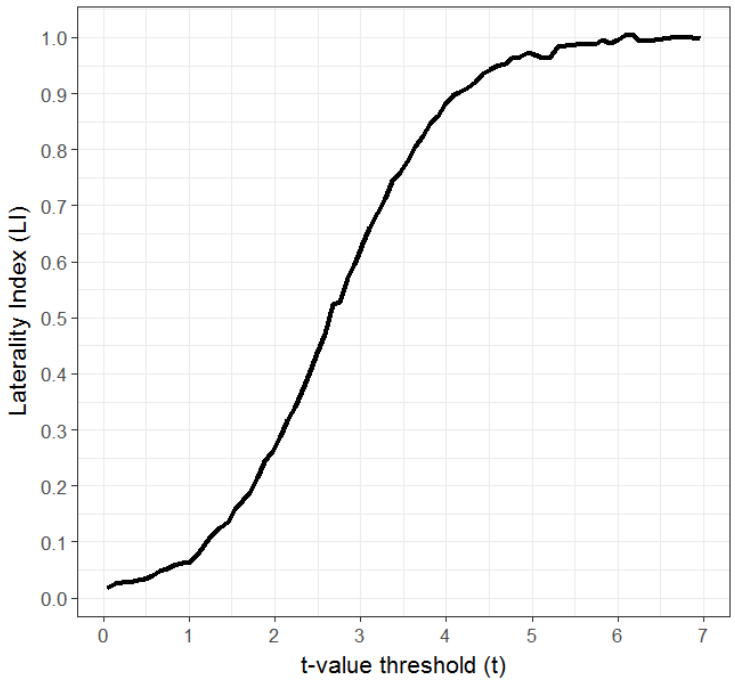


Figure 2.1: Threshold dependent laterality curve*. Plot of LI as a function of threshold (t-value).*

The purpose of this chapter is to review and evaluate different protocols for fMRI measurement of language lateralisation used by studies published between 2000 and 2016. I aimed to (1) look at the methods used by different studies over this time period in order to consider whether the field is converging on common criteria for evaluating language lateralisation, and (2) consider evidence for the robustness and reliability of these different methods in order to make recommendations for future research in this field.

2.2 Materials and Methods

A protocol for this systematic review was registered on Open Science Framework and can be found at https://osf.io/hyvc4/. This chapter addresses those objectives outlined in the protocol relating to assessment of the methods used to quantify lateralisation in fMRI studies of language lateralisation. Assessment of the impact

of language task and baselining methods will be considered in a further chapter of this thesis (*Chapter 3*).

2.2.1 *Eligibility criteria*

This systematic review was conducted on studies of fMRI language lateralisation published between 2000 and 2016. Papers were selected if they met the following inclusion criteria: (1) the paper calculated and reported LIs for language using fMRI; (2) participants were healthy monolingual adults; and (3) if participants included both patients and healthy control groups, the data for controls were reported separately. Papers were excluded if: (1) they exclusively studied structural asymmetries, children or bilingualism; or (2) they used language tasks with non-European languages. The rationale for restricting the search to studies on healthy, monolingual, adult participants was to reduce heterogeneity within the study sample.

2.2.2 *Search strategy and selection process*

The search and selection process is illustrated in in Figure 2.2. I searched Web of Science for studies published between 2000 and 2016 using the following search terms: laterali* OR asymmetr* OR dominance; AND language OR reading; AND fMRI OR functional MRI OR functional magnetic resonance imaging OR functional MR OR function MRI; NOT schizophrenia; NOT development*; NOT child*; NOT bilingual*. This was last searched on 05/12/16. I then screened the titles and abstracts of the resulting 90 papers to assess their eligibility, and subsequently conducted full-text scans to determine whether the inclusion criteria were met. This process was also carried out by a second reviewer independently. Selected lists were then compared between reviewers and any discrepancies discussed and a mutual decision made. This yielded a

total of 34 papers selected from the original 90. To ensure thorough coverage of the literature, papers citing these 34 articles were searched to look for additional articles that met criteria. From this, 50 additional papers were selected, bringing the total to 84 papers. A final search to re-check all 84 papers against search criteria identified 7 ineligible papers. During the review, a further paper was judged to not meet criteria. A list of the final 76 selected papers can be found in Appendix 1 of this thesis.

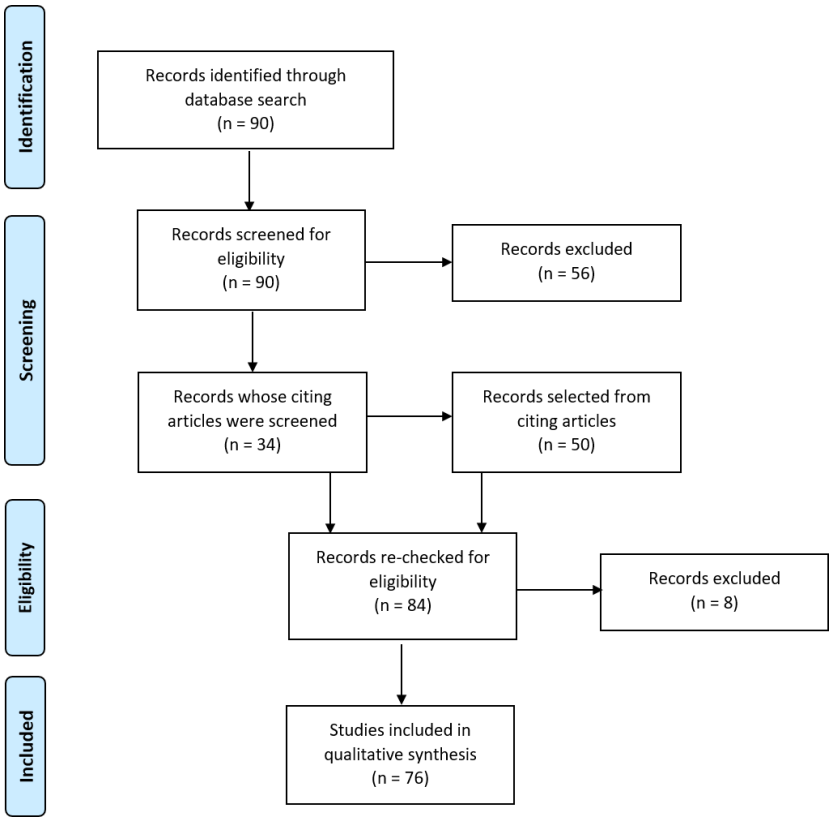


Figure 2.2: Search strategy and selection process*. Flow diagram illustrating the search and selection process for obtaining articles for inclusion in the review. Adapted from Moher, Liberati, Tetzlaff and Altman (2009).*

2.2.3 *Data collection and data summary*

For each paper, the following parameters were recorded relating to the protocol used: the type of fMRI design, the activity measures used for LI calculation, the

threshold level chosen, the use of global or regional LI calculation, the specific regions considered, the language and baseline tasks used, the use of a single or a combined task analysis and the task difficulty. Sample size and sample handedness were also recorded for each study. Information on these measures for each paper was collected and managed using REDCap electronic data capture tools (Harris et al., 2009) hosted at Oxford University. REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies, providing: 1) an intuitive interface for validated data entry; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for importing data from external sources.

The variable nature of the methods used and measures reported by different fMRI studies of language lateralisation means the data are not suitable for a meta-analysis. Instead, this chapter will document the range of methods used, and provide a qualitative summary of information from these studies that is relevant for our understanding of the robustness and reliability of LI measurement.

2.3 Results

2.3.1 *Methods for calculating LI*

As shown in Figure 2.3, a range of methods have been used to compute a laterality index in the set of studies that were considered. These will be briefly described before moving to consider the relative advantages and disadvantages of each. Note that the majority of studies used the standard LI ratio approach using the LI formula as previously outlined (55 within the search), but use of

bootstrapping approaches started to be seen in 2010 and has gained in popularity since that time. Use of different methods likely reflects their availability as well as their robustness; for example, the development of LI-tool (Wilke & Lidzba, 2007) a tool-box for LI calculation within MATLAB software (Mathworks, Natick, MA, USA) has made bootstrapping readily available and easily implementable. In contrast, other approaches such as the flip method and t-weighting may be less widely used due to the lack of dedicated software for their implementation. Relatively few studies (3) explicitly compared different methods for calculating the LI: those that did so will be covered in more detail in the following sections.

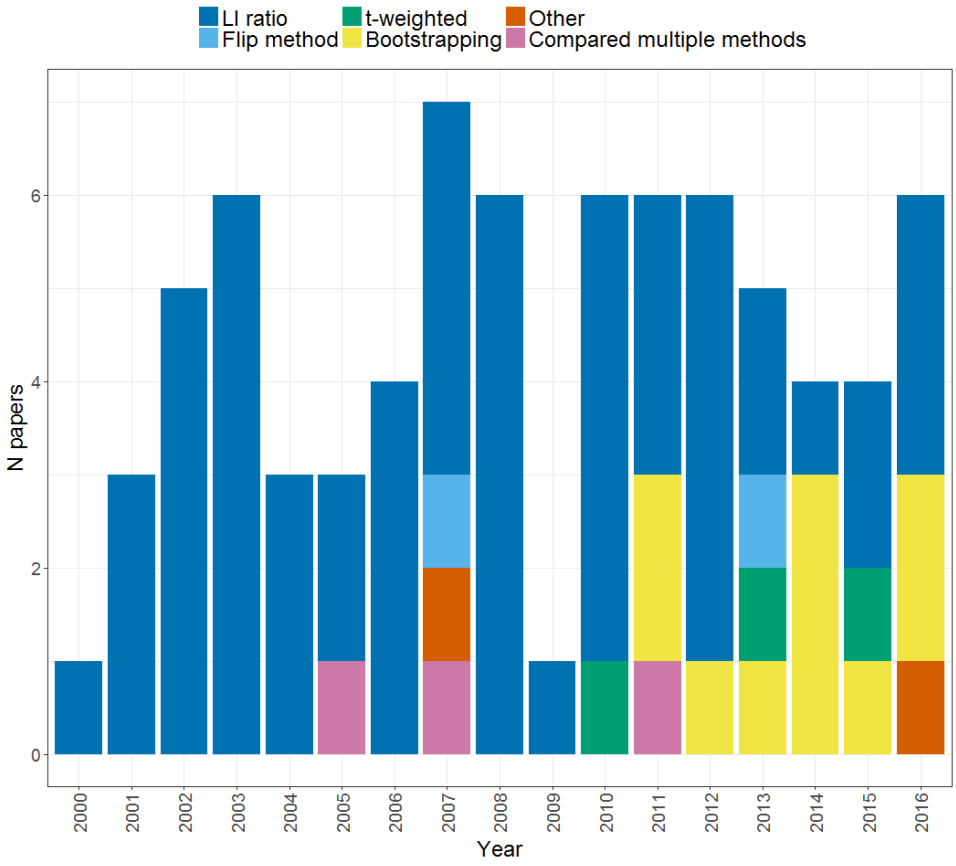


Figure 2.3: Methods of calculating an LI*. Plot shows the frequency of papers within the search using each method of LI calculation across the years from 2000 to 2016.*

2.3.2 *Thresholding*

Figure 2.4 illustrates the different thresholding approaches used by studies, and how these have changed over time. As can be seen, the majority of studies (33) used a single fixed threshold approach for determining the LI in which a single threshold level is chosen at which either extent of activation, or magnitude of activation in a given region is measured for left and right, and entered into the LI formula. As noted above however, use of a single threshold when calculating an LI is likely to yield an unreliable and inadequate measure of an individual's pattern of laterality. Awareness of this has led to a decline of the single fixed threshold approach in more recent years, in favour of approaches that aim to address the problem of threshold dependence. Each of these will be described and evaluated in turn in the following sections.

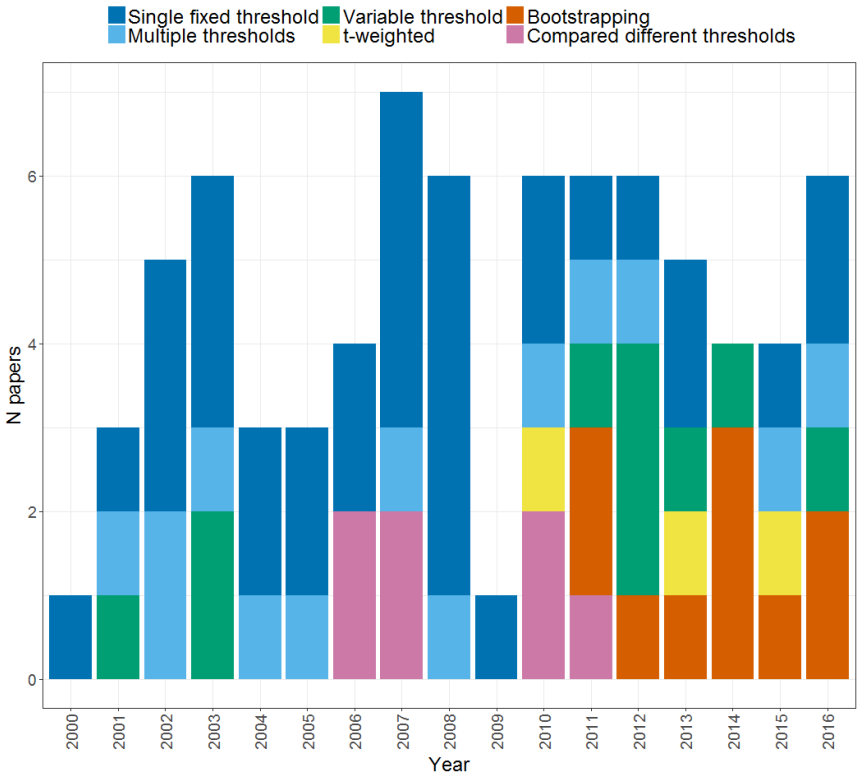


***Figure 2.4: Thresholding methods**. Plot shows the frequency of papers using each method of thresholding when calculating an LI across the years from 2000 to 2016.*

2.3.2.1 Multiple thresholds

One way to address the problem of threshold dependence is to calculate the LI across multiple thresholds. One can then produce a plot of LI as a function of threshold (see Fig. 2.1), also known as threshold dependent laterality curves (Abbott et al., 2010; Jansen et al., 2006; Ruff et al., 2008; Seghier et al., 2004; Suarez et al., 2008). Such curves can allow one to decipher an individual's general tendency towards one pattern of dominance, often showing a transition point at which the increase in laterality plateaus at a particular laterality level. However, such curves are not always informative, since in some cases they may fail to reach a plateau, or are not reproducible within a subject (Jansen et al., 2006; Rutten et al., 2002).

2.3.2.2 Variable thresholds

A second approach uses a variable or adaptive threshold, in which the threshold is set according to subject-specific parameters. One such method involves choosing that threshold which yields a fixed number of active voxels for each individual participant (Abbott et al., 2010; Fesl et al., 2010; Jansen et al., 2006; Knecht et al., 2003). Using simulated data, Abbott et al. (2010) demonstrated that thresholding at a fixed number of voxels was more robust against variability in signal strength than the standard thresholding method. They advocated plotting the LI as a function of the number of active voxels, similar to threshold dependent curves; these curves are however tighter and more stable. Furthermore, Fesl et al. (2010) reported improved reliability of LI measurement when using this variable threshold method as opposed to a single fixed threshold. However, this approach does not remove the need for arbitrary decisions, since a 'reasonable' fixed number of active voxels must be decided on. Interestingly, when using this

method Jansen et al. (2006) set the criterion number of activated voxels at a different level for each language task; this can thus enable one to take into account the fact that different tasks may require different threshold levels.

Other adaptive thresholding methods are based on setting the threshold in proportion to the maximum or mean intensity of voxels in an image. Methods include identifying the highest 5% of voxels with the highest t-values, and setting the threshold at half of their mean value (Fernandez et al., 2001; van Veelen et al., 2011); alternatively, the mean intensity of voxels within an area of interest can be used (Allendorfer et al., 2016; Partovi, Jacobi, et al., 2012; Partovi, Konrad, et al., 2012; Stippich et al., 2003; Wilke & Lidzba, 2007). Using this latter method, Wilke and Lidzba (2007) reported more stable LIs using variable threshold methods as compared to fixed thresholding, suggesting that this may make LIs more robust. This study also demonstrated a flattening of laterality curves when only those voxels that formed a significant cluster or that had a sufficiently low level of variability were included in the LI calculation. These clustering and variance weighting methods thus allow calculation of LIs to become more stable across threshold levels.

2.3.2.3 T-weighting and threshold independent methods

Alternatively, the issue of threshold dependence can be avoided by the use of a 'threshold-independent' method. One such widely used threshold-independent method is t-weighting (Branco et al., 2006; Propper et al., 2010; Suarez et al., 2008; Zaca, Jarso, & Pillai, 2013), illustrated in Figure 2.5. This involves plotting a histogram for each hemisphere of the number of active voxels against t-score threshold, and then multiplying this distribution by a weighting function that assigns weight in a way directly proportional to t-score. The integrated areas

under each hemisphere's curve can then be used as the input for the standard LI equation. Suarez et al. (2008) reported that such a method yielded reduced within-subject and between-subject variability in LI compared to fixed thresholding, resulting in clear left lateralisation across subjects.

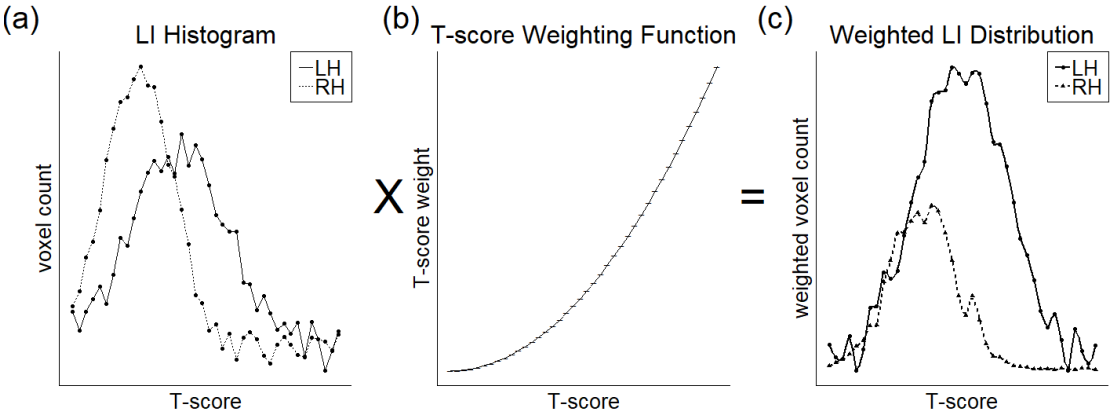


Figure 2.5: Illustration of the t-weighting method*. A plot of voxel count as a function of t-score threshold (a) is multiplied with a weighting function (b) in which higher thresholds are given greater weight, to obtain a weighted distribution (c). The integrated areas under the right and left hemisphere curves can then be used for the standard LI equation.*

Other threshold-independent methods have been developed, but these have been less widely used. Harrington, Buonocore and Farias (2006) reported that taking the average signal magnitude within activated voxels across multiple thresholds yielded higher and more reproducible LIs than a single threshold approach. Seghier, Kherif, Josse and Price (2011) used a method developed by Nagata, Uchimura, Hirakawa and Kuratsu (2001) in which the L and R terms are calculated by taking the regression of the curve obtained by plotting the number of activated voxels against threshold for each hemisphere separately. This provides a fixed term for each hemisphere that is independent of threshold for use in the LI calculation, providing a more robust measure.

2.3.2.4 Bootstrapping

A further method developed to remove the issue of threshold dependence is bootstrapping (Wilke & Schmithorst, 2006). This involves iterative resampling and calculation of LIs across multiple threshold levels, illustrated in Figure 2.6. At each threshold, a vector containing all voxel values is created from an image (b), one for each hemisphere. Multiple random samples of values (e.g. 100 samples) from these vectors are then taken (c) and an LI calculated for all possible right/left sample combinations (d). All LIs are then plotted in a histogram (e) and a trimmed mean is taken by selecting the central 50% of the data in order to reduce the effect of outliers. A weighted overall mean is then calculated from this resulting data by assigning a higher weight to higher thresholds. This method has been widely adopted in recent research on measuring language lateralisation (Berl et al., 2014; Häberling, Badzakova-Trajkov, & Corballis, 2011; Haberling, Steinemann, & Corballis, 2016; Mazoyer et al., 2014; Miro et al., 2014; Perlaki et al., 2013; Sepeta et al., 2016; Tzourio-Mazoyer et al., 2015; Van der Haegen, Cai, & Brysbaert, 2012; van der Haegen, Cai, Seurinck, & Brysbaert, 2011). As well as being threshold-independent, its key strengths include greater resistance to outliers and built-in markers for detecting the presence of outliers within the process of LI calculation (Wilke & Schmithorst, 2006) making it a robust method for assessing language laterality from fMRI data.

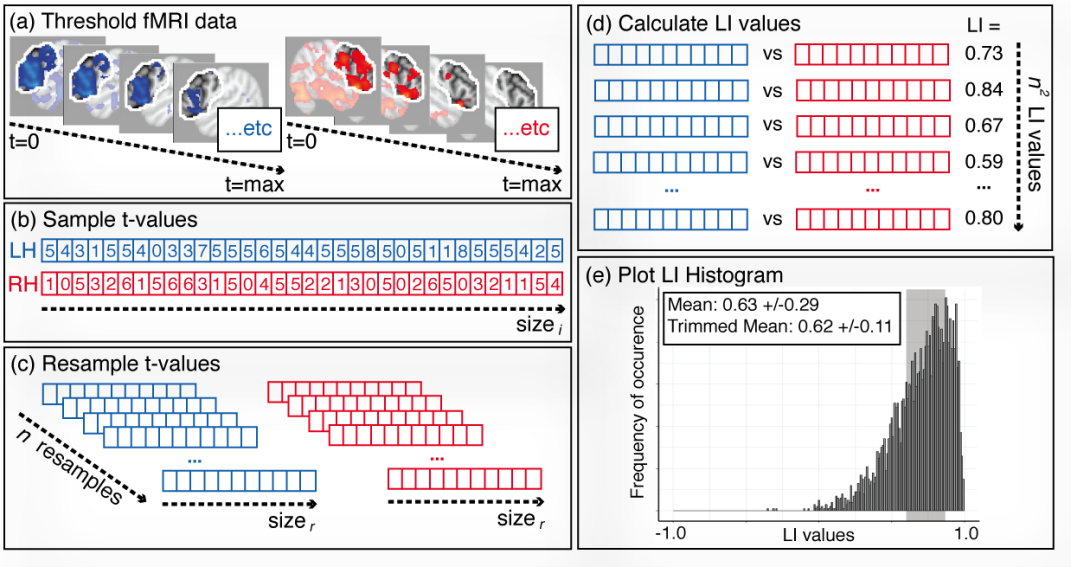


***Figure 2.6: Illustration of the bootstrapping method**. **(a) Thresholding**: contrast images are created across a range of thresholds from 0 to the maximum t-value. **(b) Sampling**: for each threshold level, a sample of t-values (size $_i$) are randomly selected from the left and right ROIs . **(c) Resampling**: values from the sample vectors are randomly resampled n times, each with size $_r$. **(d) LI calculation**: LI values are calculated for all possible combinations of right and left resamples, creating n^2 LI values in total. **(e) Histogram**: steps (b)-(c) are repeated for all threshold levels, and all of the resulting LI values are plotted in one histogram. A trimmed mean, taken from the middle 50% of the data (shaded area), is used as the final LI measure.*

2.3.3 *Activity measure*

A key decision in calculation of a laterality index from fMRI data concerns which activity measure to use; signal extent (i.e. the number of suprathreshold voxels in each hemisphere) or signal magnitude (i.e. the average intensity of suprathreshold voxels in each hemisphere). Figure 2.7 documents the different activity measures used by studies within the search. It can be seen that the majority of studies opt for an extent measure, although in recent years there has been an increased use of magnitude measures.

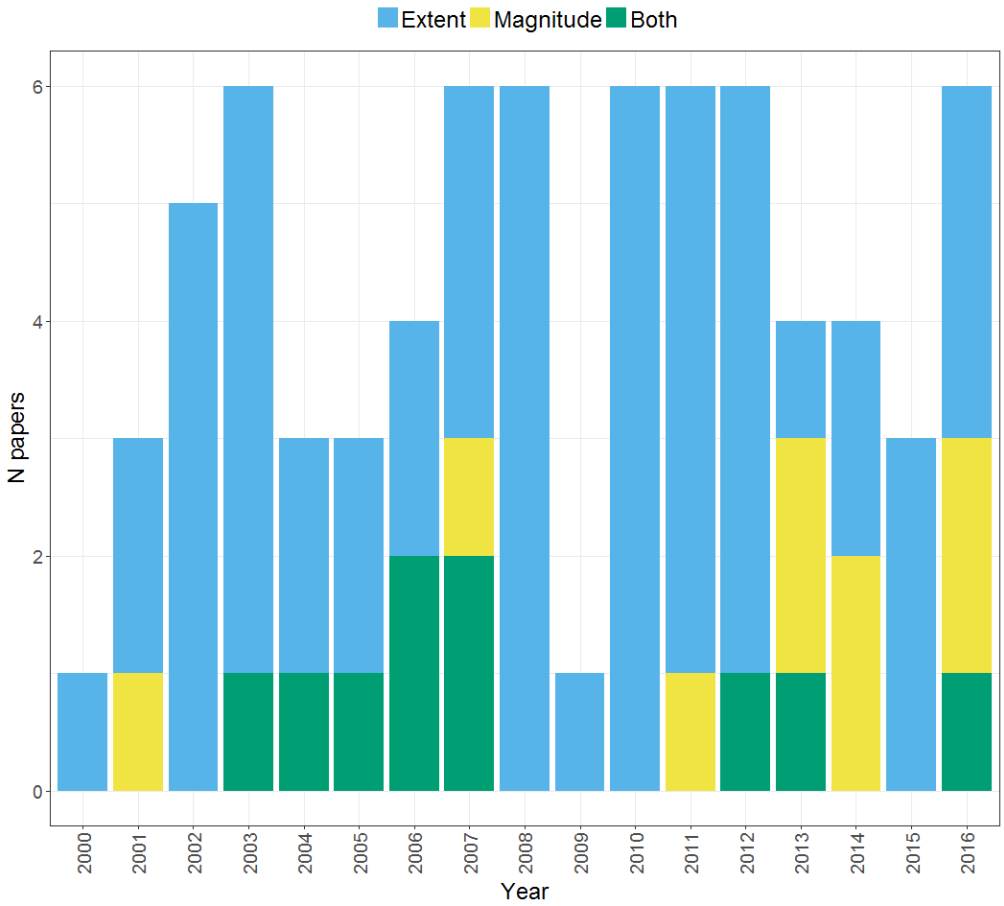


Figure 2.7: Activation measure used for LI calculation*. Plot shows the frequency of papers within the search using each type of activation measure across the years from 2000 to 2016.*

Of the studies that compare both methods, many have reported finding similar laterality indices and curves (Bethmann, Tempelmann, Bleser, Scheich, & Brechmann, 2007; Jansen et al., 2006; Ocklenburg, Hugdahl, & Westerhausen, 2013; Wilke & Lidzba, 2007). Others have reported differences in LI strength, with reports of both higher LIs for magnitude measures (Harrington et al., 2006) and higher LIs for extent measures (Jensen-Kondering, Ghobadi, Wolff, Jansen, & Ulmer, 2012). Further still, Jansen et al., (2006) reported that differences in the activity measure used for calculating laterality with a picture naming task could yield different dominance classifications for a given participant. This was not the

case however for verbal fluency and semantic decision tasks, suggesting that this may reflect something particular about the activity patterns induced by naming.

There is evidence that these measures can yield the same high levels of reproducibility for LI measurement: Morrison et al., (2016) reported 100% reproducibility for classification of language dominance and almost identical test-retest LI correlations using both activity measures for a rhyming task. However, the majority of studies report higher reproducibility for signal magnitude measures compared to signal extent measures (Adcock et al., 2003; Harrington et al., 2006; Jansen et al., 2006; Morrison et al., 2016). Importantly, Jansen et al. (2006) reported that a magnitude measure determined dominance reproducibly only when just those voxels that exceeded a criterion activation level were included; an extent measure was not reproducible. Magnitude measures have also been reported to be less sensitive to noise than a thresholded extent measure Adcock et al., (2003).

One possible explanation for higher reliability with magnitude than extent LIs relates to intersessional differences in the distribution of t-values. Gorgolewski, Storkey, Bastin and Pernet (2012) discuss the phenomenon in which shifts in t-value histograms can occur between sessions due to global effects which have a systematic effect across the whole brain. This can result in the distribution of t-values in a second session being shifted relative to the first, whilst the pattern or shape of that distribution is still preserved. In such cases, use of the same fixed threshold across both testing sessions could result in greatly different activity maps, leading to a low estimation of reliability, despite that fact that the overall pattern has not changed. Extent measures would be affected by such shifts to a

greater extent than magnitude measures, thus resulting in their reduced reliability.

Of further note is Jansen et al's (2006) finding that LIs based on signal extent lacked meaningful variation, often yielding LI values of 1; in contrast, LI magnitude measures gave greater between-subject variation in LI values. Which constitutes better laterality data depends on one's view of lateralisation measurement. That is, in cases when one wishes to classify individuals' language dominance, having LI values close to 1 or -1 would be useful to allow decisions to be clear cut. Conversely, when one is interested in quantifying individual variability in the degree of language lateralisation beyond the binary typical/atypical distinction, an LI measure that reveals greater between-subject variation would be more useful.

2.3.4 *The flip method*

The standard LI ratio does not indicate whether the difference in activity between the hemispheres is significant or not, but simply quantifies the bias in activity towards one hemisphere. The flip method (Baciu, Juphard, Cousin, & Bas, 2005) was developed to provide a direct statistical comparison of activity between the hemispheres. As illustrated in Figure 2.8, this involves contrasting two sets of functional images created for the contrast of interest (i.e. task versus control); a right side images set, in which the left hemisphere is on the left and a mirror images set, in which the image is flipped such that the left hemisphere is on the right. By contrasting these two images, one can identify those homotopic voxels that show a significant difference in activity across the hemispheres. The resulting significant voxels in each hemisphere can then be used as input for the standard LI equation.

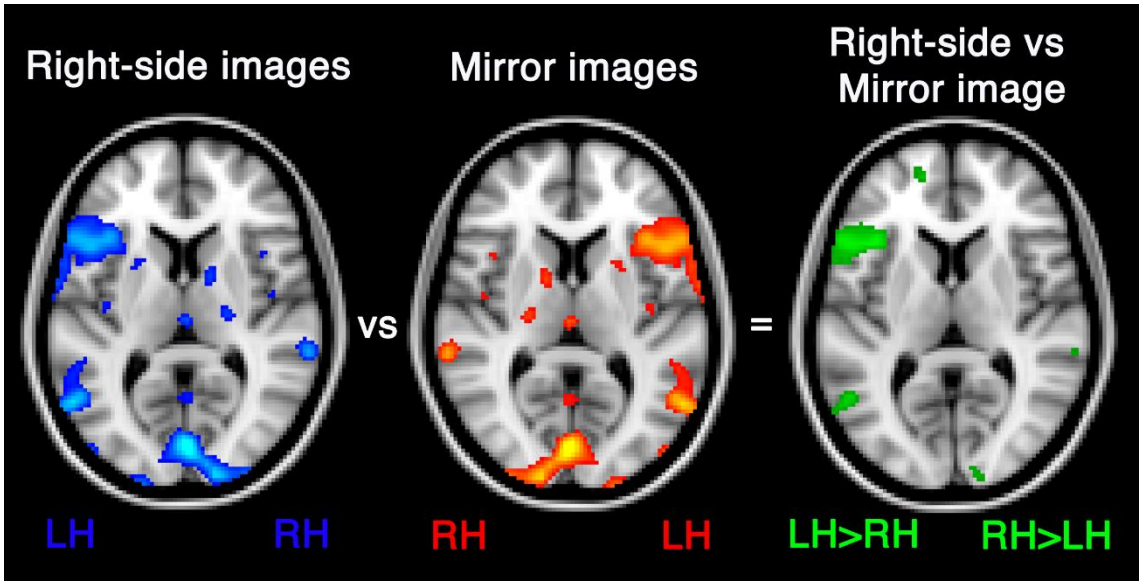


***Figure 2.8: The flip method**. By contrasting a right-side contrast image with a mirror image (flipped so that the right hemisphere is on the left), a new contrast image is generated with significant voxels indicating regions in which left activity is statistically significantly greater than right homologue activity.*

This method has been used in a small number of studies (see Fig. 2.3) as a means of measuring laterality (Cousin et al., 2007; Hernandez et al., 2013; Seghier, Kherif, et al., 2011). Baciu et al. (2005) compared the flip method to the standard LI ratio by comparing the correlations of each method's LIs with handedness lateralisation indices. While both methods yielded language LIs that were poorly correlated with handedness LIs for a verb generation task, when a rhyming task was used the flip method yielded higher correlations than the standard LI method. Such a finding is difficult to interpret, especially given the inconsistent relationship between handedness and language dominance (e.g. Groen, Whitehouse, Badcock, & Bishop, 2013).

It should be noted that the flip method cannot be used to describe bilaterality, but only non-significantly lateralised activity. Further, since significant voxels from

this contrast typically reside exclusively in the dominant hemisphere, using these voxels as input for the LI equation will often yield an LI of -/+1, rendering it an inappropriate method to provide a meaningful description of degree of lateralisation.

2.3.5 *Dominance classification*

If a categorical dominance classification is required, some form of standardised procedure is needed once an LI has been calculated. The range of methods used by different studies within the search is illustrated in Figure 2.9. The most standard method of dominance classification uses cut-offs at -0.2 and 0.2, to divide left dominance (LI > 0.2) from bilaterality (-0.2 <= LI <= 0.2) and right dominance (LI < -0.2). However, such cut-offs are arbitrary, and there were multiple studies within the search that chose their own cut-offs (see 'other cut-off values used'), including 0.1, 0.33, 0.4, 0.5 and 0.6. Thus, it can be seen that there is a high level of heterogeneity in the methods of dominance classification used by different studies. This makes it very difficult to draw conclusions and comparisons between the proportions of typically and atypically lateralised individuals reported by different studies, and thus impedes progress in understanding the distribution of such lateralisation profiles across different populations.

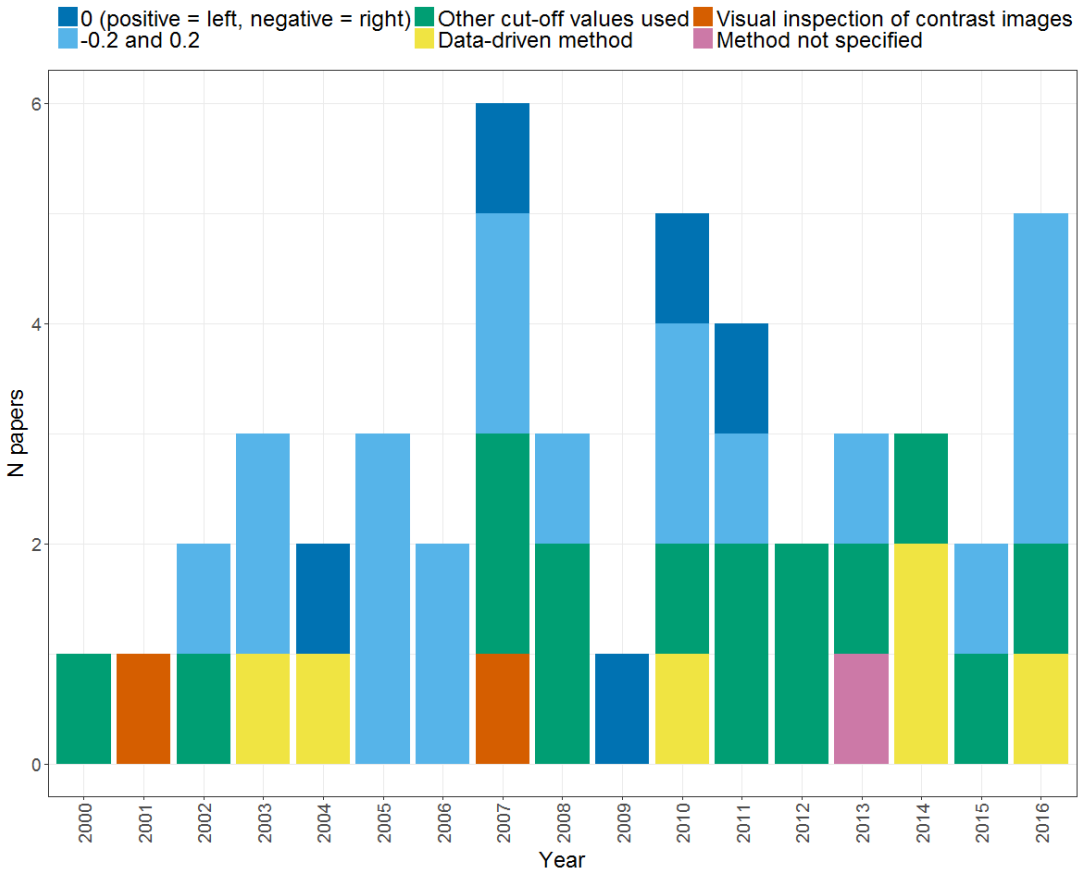


***Figure 2.9: Methods of dominance classification**. Plot shows the different methods of classifying language dominance used by studies within the search across the period from 2000 to 2016. Note that studies within the search which did not classify dominance are not included in this plot.*

Other researchers have investigated data-driven methods of defining dominance categories, in which the typicality/atypicality of cases is assessed in relation to norms within the data itself or a separate 'normative' dataset (Abbott et al., 2010; Adcock et al., 2003; Berl et al., 2014; Mazoyer et al., 2014; Seghier et al., 2004; Tzourio-Mazoyer, Joliot, Marie, & Mazoyer, 2016). One simple way of deriving cut-offs is to use 2 standard deviations below the mean as a threshold to divide typical from atypical laterality (Adcock et al., 2003; Seghier et al., 2004). Abbott et al. (2010) used a similar approach in which an individual's LI distribution (LI as a function of voxel count) was compared to a normative distribution based on a

sample of controls with 'typical' lateralisation. As illustrated in Figure 2.10, the individual's threshold-dependent laterality curve is established as being either within or below the lower 95% confidence interval for the control group (represented by the shaded area in Fig. 2.10). If an individual's laterality fell below such an interval, they were classed as atypical, on the basis of a low probability of their laterality data having come from the 'population' of the normative group. This method thus provides an objective means of establishing an individual as either atypical or typical in relation to a normative group.

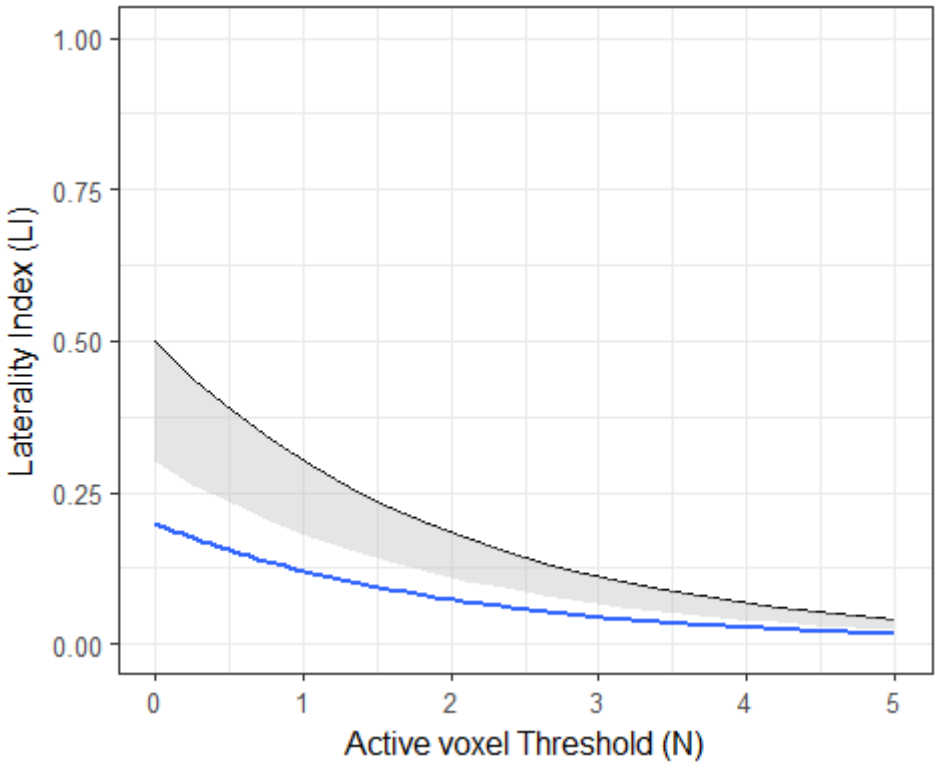


Figure 2.10: Abbott et al's (2010) method of dominance classification*. The laterality curve of the subject (blue) is compared to that of a normative control group (black), using the lower 95% confidence interval for the control group (represented by the shaded area).*

Other data-driven approaches to dominance classification suggest the existence of more than two dominance categories. Berl et al., (2014) used a hierarchical clustering method which gradually and iteratively combined cases into clusters and indicated at what level in the hierarchy the optimal cluster solution was obtained. They discussed both a three cluster solution and a two cluster solution for a large sample of right handers; the former divided subjects into left dominant, crossed dominance (one outlying case) or bilateral, whereas the latter divided typically and atypically lateralised subjects. This latter solution was found to divide at an LI of 0.5, which was therefore argued to represent a meaningful cut-off for dominance classification.

A larger-scale study by Mazoyer et al., (2014) looked at dominance categories in both left and right handed participant groups. They used Gaussian mixture modelling to extract dominance categories from laterality data (consisting of LI values between -100 and +100), which involves determining the optimal number of Gaussian functions that can be fitted to the data. This found different model solutions for right and left handed groups. For the right handed group (Fig. 2.11a), a three function solution was optimal, consisting of two overlapping 'typical' (left dominant) functions (both with LI values above +18) and a third 'ambilateral' function (consisting of LI values between -50 and +18). This agrees with Berl et al.'s (2014) study, which found no evidence for right hemisphere dominance in a right handed group. However, in the left handed group (Fig. 2.11b), an additional 'strongly atypical' function was found with strongly negative LI values (below -50). Only 10 left handers from this large sample (297 right and left handers) were strongly atypical, indicating that it is very rare. They thus argued for the need to treat atypical laterality as a heterogeneous group consisting of the subgroups of

'ambilateral' and 'strongly atypical'; conversely, it was argued that the overlapping typical distributions (from both left and right handed participants) could be combined into a single homogenous typical group.

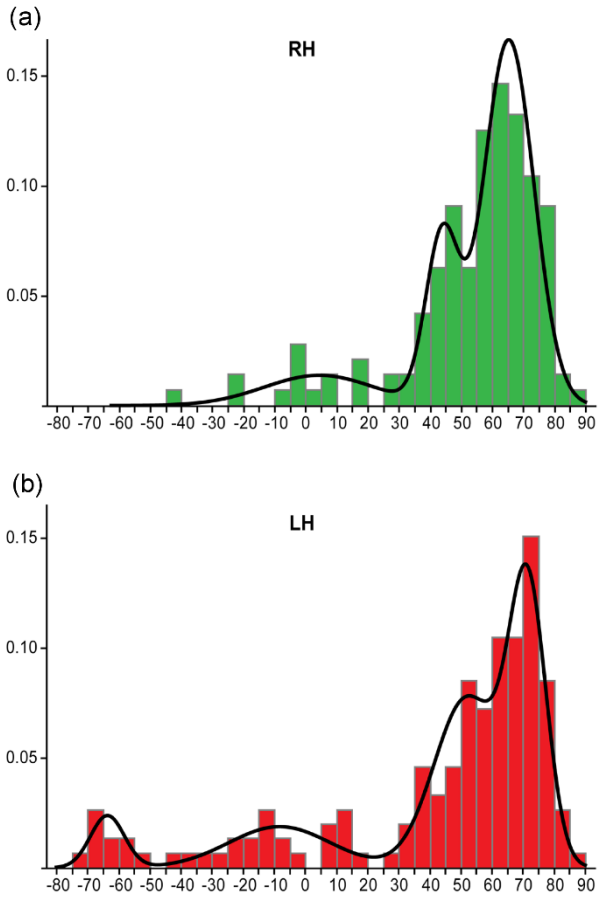


Figure 2.11: Mazoyer et al's (2014) method of dominance classification. *Histograms showing the distribution of LI values across samples of right handed (a) and left handed (b) individuals, with the envelope showing the Gaussian functions fitted to the data for determination of dominance groups. Reprinted from Mazoyer et al. (2014), open access.*

These data-driven methods of deriving dominance categories suggest that the traditional -0.2 and 0.2 cut-offs do not reflect the true distribution of LIs across individuals, and that revision of standards for dominance classification is needed in the field, with ramifications for both clinical practice and research. It should be further noted that classification of bilaterality from fMRI data is typically not reproducible across repeated testing sessions or even across different methods of LI calculation (Jansen et al., 2006), suggesting the need to interpret bilateral classification with caution. Further work is needed to establish the reproducibility of classification into the 'ambilateral' group identified by Mazoyer et al., (2014).

2.3.6 ***Effect of region of interest on laterality***

Another key consideration in calculating an LI concerns whether to include all voxels across a hemisphere (to calculate a global LI) or whether to define a region or several regions of interest (a regional LI). Figure 2.12 shows changes in approach to choice of region over time across the studies within the search. It can be seen that the majority of studies use a regional LI, either picking a single ROI or calculating multiple LIs from different regions. A number of studies have compared global and regional approaches to LI calculation, particularly in earlier years within the search period. This evidence on the relative advantages and disadvantages of the global and regional approaches is discussed below. It can also be seen that a substantial proportion of studies use a single ROI in LI calculation; however, a review of evidence on regional variability in laterality questions the adequacy of such an approach (see *2.3.6.3 Regional heterogeneity in laterality*).

2.3.6.1 Global and regional LIs

A potential issue with global LIs is that voxels outside the areas most relevant to the language paradigm can have a strong influence on the LI obtained. Whilst many studies have shown that LIs from pre-specified language ROIs are stronger and more reliable than global LI measurements (Fernandez et al., 2001; Pravata et al., 2011; Rutten et al., 2002; Suarez et al., 2008), others have found no difference (Hund-Georgiadis, Lex, Friederici, & von Cramon, 2002) or the opposite (Rutten et al., 2002; Wilke & Lidzba, 2007). Further studies suggest that whether global or regional LI measures are optimal may depend on

other methodological decisions. For example, Jansen et al., (2006) reported that global LIs were more reliable than regional LIs when a voxel count approach was

used. Similarly, Rutten et al., (2002) reported greater reliability for global LIs over regional LIs during an antonym generation task, but that the reverse was true for a verb generation task.

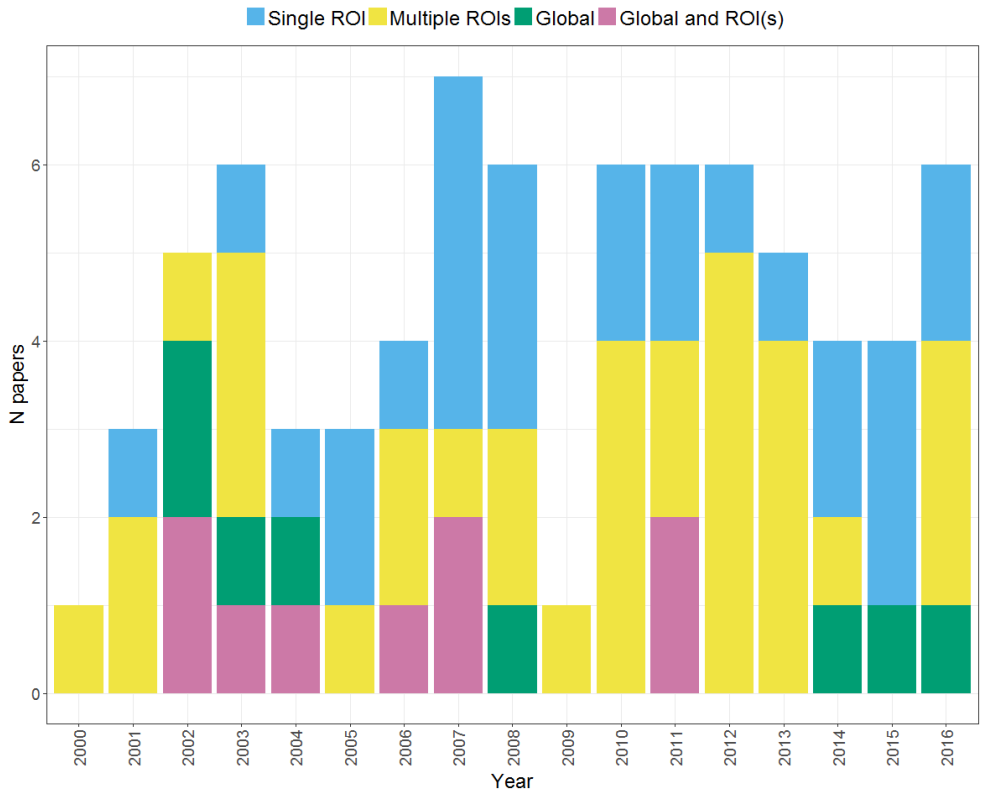


***Figure 2.12: Use of global and regional approaches to LI calculation over time**. Plot shows the regional approaches used for LI calculation by studies within the search across the period from 2000 to 2016.*

In general however, cases of crossed dominance or regional heterogeneity in lateralisation have been used to argue for the need for regional rather than global laterality indices, which fail to capture these finer grained individual patterns in regional laterality (e.g. Seghier et al., 2011). Such evidence will be discussed in the following section, *Regional heterogeneity in laterality*.

2.3.6.2 Regional approaches to LI calculation

When using a regional rather than a global LI, careful thought must be given to deciding which areas to choose as ROIs. Typically, areas within frontal and temporoparietal cortex are chosen for measuring language laterality, such as the inferior frontal cortex or posterior superior temporal gyrus. It is not the case that any single ROI or combination of ROIs will always be optimal for assessing laterality; instead, the choice of ROI(s) must be guided by other factors such as the language function being studied or the purpose of laterality measurement.

There are a number of studies which measure laterality from both frontal and temporoparietal ROIs, which allows one to compare the robustness and reliability of their LIs. There is mixed evidence over whether frontal or temporoparietal ROIs yield stronger or more reliable LI values. The majority of studies report stronger laterality in frontal than temporoparietal ROIs, across a wide range of both expressive and receptive tasks (Clements et al., 2006; Gaillard et al., 2003; Harrington et al., 2006; Niskanen et al., 2012; Ocklenburg et al., 2013; Partovi, Jacobi, et al., 2012; Partovi, Konrad, et al., 2012; Propper et al., 2010; Szaflarski et al., 2008; Vernooij et al., 2007; Vikingstad, George, Johnson, & Cao, 2000). However, some have reported the opposite, a pattern particularly associated with the use of receptive tasks such as semantic decision, speech listening or auditory comprehension (Bethmann et al., 2007; Brennan et al., 2007; Fernandez et al., 2001; Häberling et al., 2016; Harrington et al., 2006; Hund-Georgiadis et al., 2002; Jensen-Kondering et al., 2012; Niskanen et al., 2012; Ramsey, Sommer, Rutten, & Kahn, 2001; Sanjuan, Forn, et al., 2010a; van Oers et al., 2010). A similar pattern emerges for reliability of laterality measurement; while frontal LIs are often reported as more reliable than temporoparietal LIs (Harrington et al.,

2006; Partovi, Konrad, et al., 2012; Szaflarski et al., 2008), the reverse can be true when using a receptive language task (Harrington et al., 2006; Jansen et al., 2006). This suggests that the two areas are both capable of yielding robust and reliable laterality measurement, provided that a receptive task is used to engage temporoparietal areas.

2.3.6.3 Regional heterogeneity in laterality

Research comparing laterality across regions has reported cases of crossed or dissociated dominance across different cortical language areas, at both an individual and a group level (Berl et al., 2014; Bethmann et al., 2007; Häberling et al., 2016; Jansen et al., 2006; Propper et al., 2010; Seghier, Kherif, et al., 2011; Thivard et al., 2005; van der Haegen et al., 2012; Vikingstad et al., 2000; Vingerhoets et al., 2013). Bethmann et al., (2007) reported four subjects with crossed frontal and temporal dominance for a semantic decision task; in particular, one subject was classified as bilateral when ROIs were combined, whereas classification based on either only a frontal or a temporal ROI yielded contradicting dominance categories. Other studies report within-subject dissociations between the dominance measured from different regions using different language tasks. Cases of crossed dominance have been reported between frontal LIs for expressive language and temporal or occipitotemporal LIs for receptive language (Häberling et al., 2016; van der Haegen et al., 2012), and between temporal and frontal LIs for two different expressive language tasks (Vikingstad et al., 2000). In Van der Haegen et al.'s (2012) study, the majority of left handed subjects showed colateralisation of inferior frontal cortex for verb generation and ventral occipitotemporal cortex for lexical decision, but a small minority (3 out of 57 participants) showed crossed dominance. Such crossed

dominance is not limited to left handers however; Häberling et al., (2016) reported cases of crossed frontal expressive and temporal receptive dominance among both left and right handers.

Such cases highlight the inadequacy of relying on a single global or regional LI, and how this practice in previous research may have led to an underappreciation of such crossed dominance in the literature. Indeed, such regional variation is not only found at the individual level, but also at a group level. In a comparison of lateralisation for a semantic word matching task across 50 different ROIs, Seghier et al., (2011) reported a negative correlation between LIs obtained from the angular gyrus and the ventral precentral gyrus. This suggests that regional heterogeneity can be a normal part of typical profiles of hemispheric lateralisation for language.

2.4 Summary and conclusions

This chapter has highlighted the many different ways in which calculation of an LI from fMRI data presents a methodological challenge, and how the use of different methods has changed over time. Of course, the decisions one makes when designing an fMRI laterality experiment will depend on the question being investigated; however, this review has highlighted some key principles that emerge from the literature that should be considered in order to generate increased standardisation in fMRI laterality protocols across future studies. Increased homogeny in the methods used by different studies will enable better integration of research findings in order to draw conclusions as to the nature and correlates of language lateralisation.

FMRI LI calculation must address the problem of threshold dependence. Bootstrapping represents a promising method for calculating a robust, threshold-independent LI, making it a widely used method in recent research. The general pattern of evidence suggests that signal magnitude may provide a more robust and reliable measure than signal extent, and that regional LIs calculated from pre-specified ROIs are stronger and more reliable than global LIs. However, such decisions need to be considered in light of other methodological parameters (e.g. the activity measure used) in order to optimise the fMRI analysis. A useful tool for implementing such analysis methods is LI-tool, a tool-box within MATLAB software (Mathworks, Natick, MA, USA), developed by Wilke and Lidzba (2007). This includes options for different thresholding techniques and activity measures, and can implement the bootstrapping method.

Data-driven methods can provide a less arbitrary means of classifying language dominance and support the validity of a three category model of language dominance within a mixed handedness sample, consisting of typical (left dominant), ambilateral, and atypical (strongly right dominant) groups; conversely in right handed samples a two-category model (typical versus ambilateral) may be sufficient (Mazoyer et al., 2014). No subsequent studies have implemented the thresholds for dominance classification suggested by Mazoyer et al's (2014) large scale study, except a paper reporting on the same sample of right and left handed individuals (Tzourio-Mazoyer et al., 2016). Further work is needed to implement and validate these cut-offs, to see if these generalise to other samples and thus could be used as standard practise.

The choice of which regions of interest to use for LI calculation again depends on the question being asked. If one wishes to classify laterality for a particular

language function, one must consider which ROI yields the highest and most reliable LIs for that function. Frontal ROIs typically yield the strongest and most reliable laterality for expressive tasks, whereas a temporoparietal ROI may be more appropriate for receptive tasks. However, measurement of laterality from a single regional or global ROI can be misleading and does not capture potential regional heterogeneity. This was highlighted by Tailby, Abbott and Jackson (2017) in relation to the need to appreciate such regional variability in presurgical planning with epilepsy patients, and the consequent inadequacy of a single metric to quantify an individual's 'language dominance'. Therefore, in fMRI laterality protocols, lateralisation across frontal and temporoparietal ROIs for at least one expressive and one receptive task should be measured, to obtain a comprehensive picture of any individual's pattern of hemispheric dominance for language. This will enable further work to investigate the significance of such regional heterogeneity in dominance; for example, are there any functional consequences of having crossed frontal-temporal language laterality? In this way, fMRI as a method of laterality measurement can provide unique insights into lateralisation at a regional level; this should be fully exploited in future research.

Lastly, it should be noted that fMRI data provides opportunities for investigating the nature of lateralisation for a function that goes beyond simple calculation of laterality indices. For example, fMRI can be used to investigate more mechanistic questions such as potential relationships between inter- and intrahemispheric connectivity and lateralisation (e.g. Frässle, Krach, Paulus, & Jansen, 2016; Seghier, Josse, Leff, & Price, 2011). In order to pursue more advanced analyses, it will be important for the field to establish robust protocols for laterality measurement as standard across studies.

Chapter 3 Measuring language lateralisation with different language tasks: A systematic review

Abstract

Traditionally, lateralisation for language has been viewed as a unitary concept. Recently however, there has been increasing awareness of the possibility that different language processes may develop hemispheric lateralisation independently, and to varying degrees. However, it is not always clear whether differences in laterality across language tasks with fMRI are reflective of meaningful variation in hemispheric lateralisation, or simply of trivial methodological differences between paradigms. This chapter presents a systematic review that aims to assess different language tasks in terms of the strength, reliability and robustness of the laterality measurements they yield with fMRI, to look at variability that is both dependent and independent of aspects of study design, such as the baseline task, region of interest, and modality of the stimuli. Recommendations are made that can be used to guide task design; however, this review predominantly highlights that the current high level of methodological variability in language paradigms prevents conclusions as to how different language functions may lateralise independently. Suggestions are made for future research using tasks that engage distinct aspects of language functioning, whilst being closely matched on non-linguistic aspects of task design (e.g. stimuli, task timings etc); such research could produce more reliable and conclusive insights into language lateralisation. This systematic review was

registered as a protocol on Open Science Framework (https://osf.io/5vmpt/), and has been published as a paper (Bradshaw, Thompson, Wilson, Bishop, & Woodhead, 2017).

3.1 Introduction

Although there has been a body of work considering between individual variability in lateralisation for language, less research has considered the extent and potential significance of within-individual variability in laterality. Because both expressive and receptive aphasia are so reliably associated with left-hemisphere injury, there tends to be an assumption that left-sided lateralisation is a general feature of language processing, consistent across language domains. Nevertheless, there is evidence that lateralisation may differ within individuals for different language functions, as well as between individuals in side and extent.

Early suggestions of within-individual variability can be found in Rasmussen and Milner's (1975) reports of Wada testing in patients undergoing surgery to treat epilepsy. They reported on several patients with bilateral speech representation, manifest as a dissociation between the hemispheric organisation of different language functions. Specifically, while anaesthetic injection to one hemisphere selectively disrupted naming and not verbal serial order tasks (e.g. reciting the days of the week), an injection to the other hemisphere produced the reverse pattern. This was construed as evidence that in some cases a 'division of labour' can exist between the hemispheres, in which different 'speech centres' can lateralise to different hemispheres independently. Although such evidence was from the study of a special population, it was argued that such a phenomenon should not necessarily be considered as a result of the type of brain damage and reorganisation that occurs in epilepsy. This raises the possibility that cerebral

lateralisation may be a multifactorial rather than a unitary process, with different language processes developing hemispheric lateralisation independently, and to varying degrees (Bishop, 2013).

Indeed, several contemporary models of language predict different patterns of lateralisation for different language processes (Hickok & Poeppel, 2007; Peelle, 2012; Poeppel, 2014; Price, 2012). These predictions are summarised in Table 3.1. Different models make different distinctions between language processes and use different terminology, but some general patterns emerge. Acoustic processing of speech input and speech articulation are generally considered to be bilateral, whereas comprehension and generation of more meaningful language is considered to be lateralised. There are some points of disagreement between theories however, either in terms of the extent of lateralisation for a particular language process or the theoretical reasons proposed for such patterns of lateralisation.

Contemporary non-invasive techniques allow more extensive research on patterns of laterality than earlier clinical studies. Functional magnetic resonance imaging (fMRI) has been used to investigate lateralisation using a range of different language tasks. At first glance, the literature appears to support the notion that language laterality is not unitary, because we can see differences between tasks in the strength of the laterality measurements they yield. However, the reasons for such variability in LI strength across language tasks can be debated; could it simply be an artefact of more trivial differences in task design, or does it reveal something fundamental about the hemispheric organisation of different components of language? Completely equating tasks that engage different language functions is a rather unrealistic goal; for example, how does

one devise tasks so as to equate and compare two such diverse language functions as speech production and speed comprehension? However, more attempt is needed in research to optimize paradigms for LI measurement, in order to try to reduce the possibility of differences in task sensitivity or measurement error being responsible for variability in LIs across tasks.

Table 3.1: Model-based predictions of language lateralisation. *Table illustrates some predictions of different models of the neural basis of language, in terms of the lateralisation expected for different aspects of language processing. L = lateralised, B = bilateral.*

Theoretical principle/ Model	***Speech acoustic processing***	***Speech comprehension***	***Speech articulation***	***Semantics***	***Syntax***
Dual stream model of speech processing (Hickok & Poeppel., 2007)	B	B	L	B	-
Hierarchical asymmetry of linguistic complexity (Peelle., 2012)	B	L	-	L	L
Bilateral sensorimotor inputs/outputs and left lateralised central language processes (Price., 2012)	B	L	B	L	L
COM-PRE hypothesis (Poeppel., 2014)	B	L	B	L	L

This chapter aims to assess and evaluate evidence on the robustness of laterality measured using fMRI with different language tasks, from studies published between 2000 and 2016. This was done with a view to providing some guidance

on optimizing variables such as region of interest and baseline task on a task-by-task basis. Such optimization will be important before fMRI tasks can be used to systematically probe patterns of co-lateralisation and independent lateralisation of different language functions. Further, this review was conducted with a view to providing evidence-based insights into optimal task design that could be used to guide development of my own tasks for laterality measurement with fTCD; these were used in a study presented in Chapter 5 of this thesis. Overall, I hypothesised that 1) different language tasks will demonstrate different levels of lateralisation and 2) parameters such as the region of interest and baseline task used will have effects on laterality measurement that may be task-specific.

3.2 Materials and Methods

The literature search methodology and set of selected papers for the systematic review presented in this chapter is identical to those used in Chapter two; the reader is thus referred back to this chapter (see *Chapter 2: 2.2 Materials and* method) for information on the search strategy and selection process.

3.2.1 *Data analysis*

The variable nature of the methods and measures reported by different papers did not permit performance of a meta-analysis. Instead, to illustrate the strength of laterality measured across different language tasks, forest plots were produced showing the mean and 95% confidence intervals of LI values reported in the studies, as well as their associated methods of LI calculation, region(s) of interest, and language and baseline tasks (Fig. 3.1, 3.2 and 3.3). These outcome measures were not always available in every paper however; LI values and/or their spread were sometimes omitted altogether or given in a different form e.g. median values. Where standard deviation or standard error were given, these

were converted to 95% confidence intervals. Forest plots were created using a script in R, which is available along with the data on open science framework (https://osf.io/7s4hv/).

To avoid the potential confound of heterogeneity in samples in terms of handedness, forest plots included only mean LIs that were reported either from solely right handed participants, or from mixed handedness samples where the relative proportion of left and right handers was representative of the general population (around 10% left handed, 90% right handed). LIs reported from studies that selected a participant group on the basis of their pre-known lateralisation were excluded. Where more than one frontal LI was reported from a study, IFG LIs were selected; where more than one temporoparietal LI was reported, that LI calculated from the largest area of temporoparietal cortex was selected.

3.3 Results

The main language tasks identified in the literature search are listed in Table 3.2, with counts of the number of studies using each task (one study is missing from these counts as their language task did not fit in to any of these categories). Mean LIs reported from studies using these different language tasks are given in Figures 3.1, 3.2 and 3.3. A single language task typically engages multiple language processes in an overlapping fashion. This may either be because of task requirements, or reflect spontaneous engagement of task irrelevant processing by the perception of linguistic stimuli. Table 3.2 provides one characterisation of the different language processes engaged by each of the language and baseline tasks included in this review. Comparing the language processes engaged by active and baseline tasks is crucial, because when activation for a language task is subtracted from a baseline the aim is to isolate

specific linguistic functions, and the extent to which this is successful will depend on the demands of the baseline task.

The following review will discuss each language task in turn, with reference to the involvement of different language processes and the forest plots of mean LI values (Fig. 3.1, 3.2 and 3.3). Table 3.2 highlights the difficulty in designing a task which isolates a single language function in order to study its laterality; this must be kept in mind when interpreting LI values and theorising on the lateralisation of particular language processes.

Table 3.2 (on next page): Language and domain general processes engaged by different language and baseline tasks. *Table shows the main language tasks (top left quadrant) and baseline tasks (bottom left quadrant) identified as being widely used in laterality research. For each type of task, the number of studies (N) within the current literature search using this task is given, as well as one characterisation of the different language processes (middle column) and domain general processes (right column) they engage.* ✓ *= engaged,* **(✓)** *= sometimes engaged (e.g. depending on task demands, modality of stimuli, occurrence of automatic linguistic processing).*

		Language processes						Domain general processes		
Task	N studies	Speech motor planning/ articulation	Phono-logical access	Ortho-graphical processing	Semantics	Word retrieval	Syntax	Working memory	Motor processing	Auditory processing
Language tasks										
Verbal fluency	53	✓	✓	(✓)	(✓)	✓		✓		
Sentence generation	5	✓	✓		✓	✓	✓	✓		
Passive speech listening	7		✓		✓		✓	(✓)		✓
Text reading	2	✓	✓	✓	✓		✓	(✓)		
Phonemic decision	8	(✓)	✓	(✓)	(✓)	(✓)		✓	(✓)	
Semantic decision	20		(✓)	(✓)	✓	(✓)		✓	(✓)	
Sentence comprehension	8		✓	(✓)	✓		✓	✓		(✓)
Naming	9	✓	✓		✓	✓				
Baseline tasks										
Rest	37				(✓)					
Perceptual decision (non-linguistic)	25							✓	(✓)	
Finger tapping	3								✓	
Non-word/word repetition	8	✓	✓							
Recite months of the year/count sequence	2	✓	✓		(✓)	✓				
Tone listening	3									✓
Backward speech listening	2									✓
Nonsense text reading	2	✓	✓	✓						

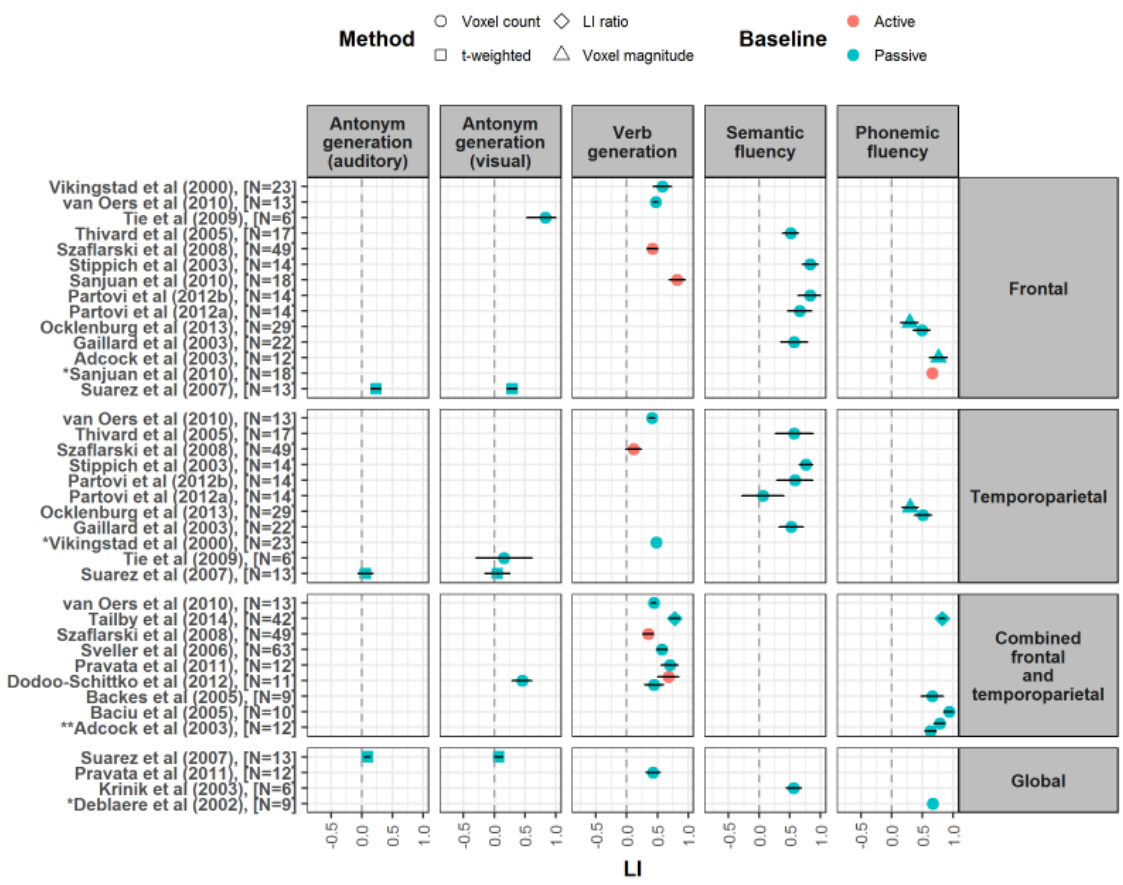


Figure 3.1: Forest plot of mean LI values for verbal fluency tasks. *Plot is divided up according to region of interest used for LI calculation (frontal, temporoparietal, combined frontal and temporoparietal, and global). Error bars represent 95% confidence limits. Colour of symbol indicates type of baseline task used (active or passive), and shape of symbol indicates method of LI calculation (see key). *Papers did not report a measure of spread for LI values, so confidence interval is not shown. **LI values reported by this paper are given at different thresholds: Z = 5.3 (top), Z = 2.3 (bottom).*

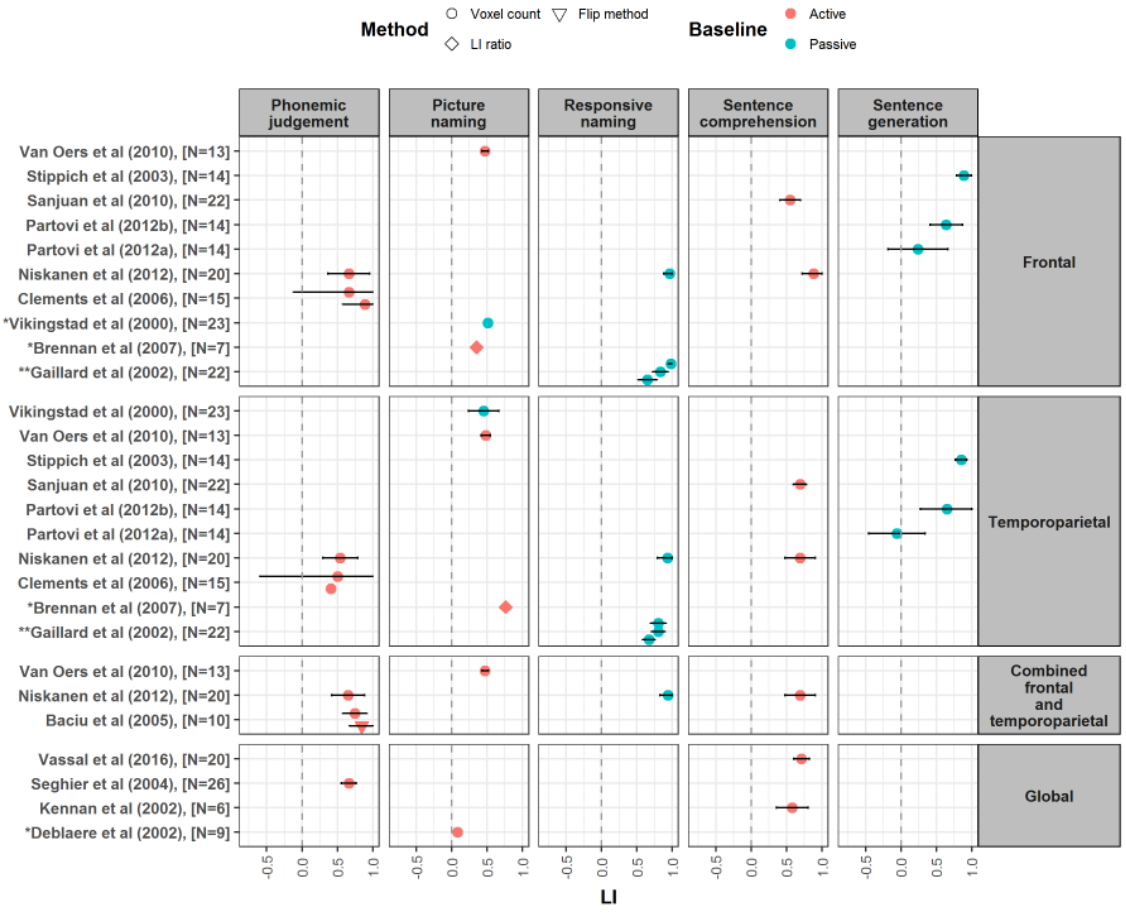


Figure 3.2: Forest plot of mean LI values for phonemic judgement, naming, sentence comprehension and sentence generation tasks. *See caption for Figure 3.1 for detailed description. *Papers did not report a measure of spread for LI values, so confidence interval is not shown. **LI values reported by this paper are given at different thresholds: t = 5 (top), t = 4 (middle) and t = 3 (bottom).*

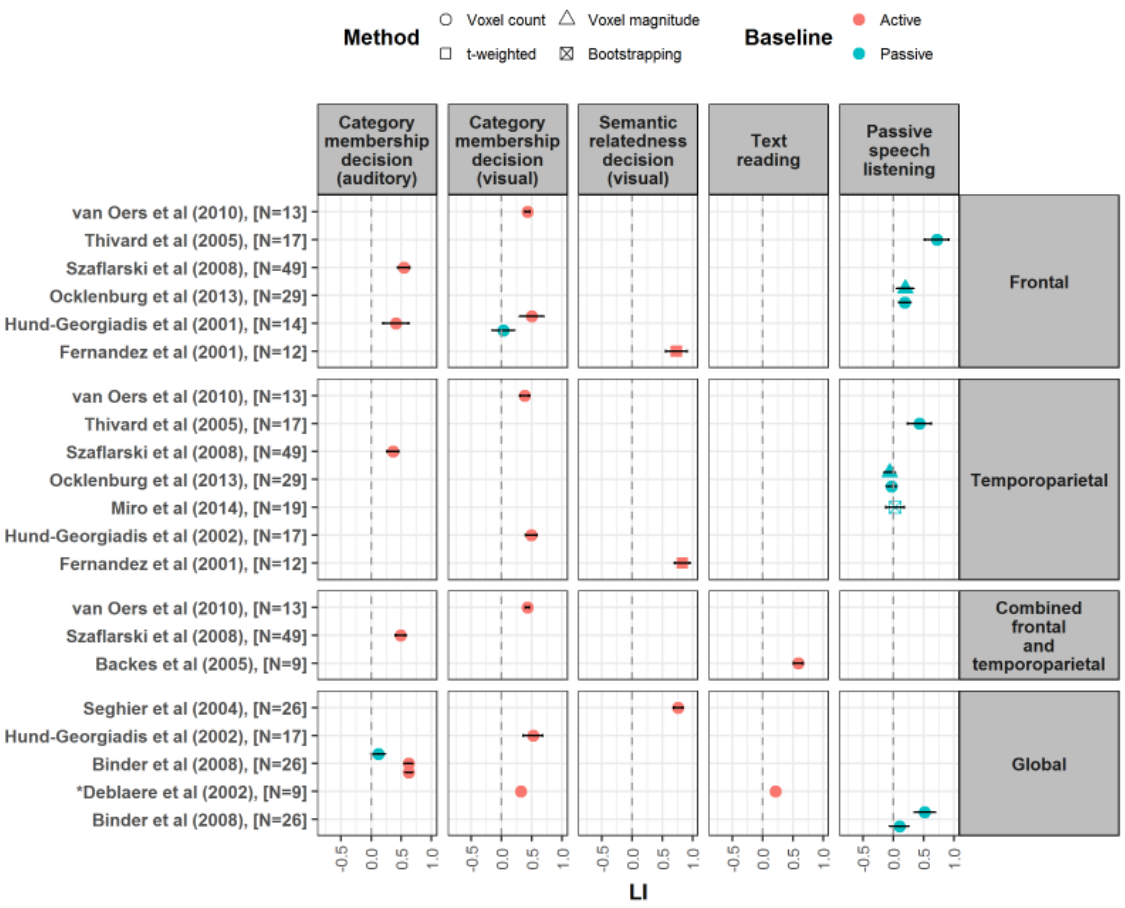


Figure 3.3: Forest plot of mean LI values for semantic decision, text reading and speech listening tasks. *See caption for Figure 3.1 for detailed description. *Papers did not report a measure of spread for LI values, so confidence interval is not shown.*

3.3.1 *Verbal fluency*

Verbal fluency tasks have traditionally been viewed as the gold standard for measuring language lateralisation with fMRI. Here, the participant must generate (covertly or overtly) words that meet certain criteria, such as beginning with a particular letter (phonemic fluency), belonging to a particular semantic category (semantic fluency), verbs that are semantically associated with a particular noun (verb generation), or words that are antonyms/synonyms (antonym/synonym generation). Any lateralisation induced by this task may thus reflect a mixture of phonological, semantic, word retrieval and speech motor planning/articulation processes (see Table 3.2). Lateralisation of speech motor processes is a subject of debate (see Table 3.1), with some considering them left lateralised (Hickok & Poeppel, 2007), but others bilateral (Poeppel, 2014; Price, 2012).

3.3.1.1 LI strength, reproducibility and robustness

Across the papers reviewed here, verbal fluency tasks are consistently reported as yielding the strongest laterality when compared to other receptive and expressive tasks within studies (Baciu, Watson, et al., 2005; Deblaere et al., 2002; Dodoo-Schittko, Rosengarth, Doenitz, & Greenlee, 2012; Harrington et al., 2006; Jensen-Kondering et al., 2012; Niskanen et al., 2012; Ocklenburg et al., 2013; Ramsey et al., 2001; van der Haegen et al., 2012; Vikingstad et al., 2000; Zaca et al., 2013). Studies included in the forest plots produced here report a wide spread of LI values for verbal fluency tasks, ranging from 0.05 to 0.94 (see Fig. 3.1).

A number of studies have compared the use of different verbal fluency paradigms. When a frontal ROI is used, semantic fluency is reported as less strongly lateralising than verb generation or phonemic fluency (Kleinhans,

Mueller, Cohen, & Courchesne, 2008; Ruff et al., 2008; Sanjuan, Bustamante, et al., 2010). Conversely, two studies using combined frontal and temporoparietal ROIs reported no differences in their strength of laterality (Rutten et al., 2002; Tailby, Weintrob, Saling, Fitzgerald, & Jackson, 2014). This suggests that these tasks may differ in the extent of lateralisation they induce across different language areas (see *3.3.1.2 Effect of region of interest*). Interestingly, multiple studies report finding that LIs from generation tasks can vary substantially depending on methodological choices made when calculating laterality, such as the threshold chosen (Dodoo-Schittko et al., 2012), the use of normalisation, smoothing and clustering techniques (Baciu, Watson, et al., 2005), and the activity measure used (Harrington et al., 2006).

3.3.1.2 Effect of region of interest

Verbal fluency tasks tend to induce the strongest laterality in frontal ROIs (Gaillard et al., 2003; Niskanen et al., 2012; Ocklenburg et al., 2013; Partovi et al., 2012a; Partovi et al., 2012b; Propper et al., 2010; Szaflarski et al., 2008; Vernooij et al., 2007; Vikingstad et al., 2000; Vingerhoets et al., 2013; Zaca, Jarso, & Pillai, 2013). Although they can induce strong laterality in temporoparietal ROIs (Harrington et al., 2006; Jensen-Kondering et al., 2012; Stippich et al., 2003), this may not be significantly greater than other tasks within this ROI (Zaca et al., 2013). However, this may depend on the particular fluency task used; Jensen-Kondering et al. (2012) reported that while phonemic fluency and verb generation yielded the strongest lateralisation for a frontal ROI, the strongest laterality for a temporoparietal ROI was seen with semantic fluency, consistent with a role for such areas in semantic cognition.

3.3.1.3 Effect of baseline task

Whilst the majority of studies using verbal fluency tasks tend to employ a passive baseline task such as fixation, a number have used active baselines such as finger tapping or silent word repetition (see Fig. 3.1 and Table 3.2). Dodoo-Schittko et al. (2012) reported that an active baseline task which required subvocal manipulation of the order of syllables within a pseudoword yielded significantly stronger laterality for a verb generation task compared to the use of a passive resting baseline. This is consistent with the idea of subtraction of bilateral speech motor planning related activity (Poeppel, 2014; Price, 2012).

3.3.2 ***Sentence generation***

Sentence generation requires participants to generate sentences to describe presented pictures. These sentences may either be pre-defined and learnt prior to scanning, or generated during the scan itself. Sentence as opposed to word generation involves additional syntactic and semantic-integration processes relating to the construction of a sentence (see Table 3.2). These are argued to be left lateralised by multiple models (Peelle, 2012; Poeppel, 2014). Poeppel's (2014) COM-PRE hypothesis makes a distinction between bilateral processing within input and output interfaces (e.g. auditory perception and speech production), and left dominant processing of combinatorics and composition (COM) or linguistically-based predictions (PRE). Similarly, Peelle (2012) predicts that while unconnected language (e.g. single words) is processed bilaterally, processing of connected language that requires more complex linguistic operations is left lateralised. Thus, these models might predict stronger laterality for sentence generation over word generation paradigms, due to the additional sentential processing demands.

3.3.2.1 LI strength, reliability and robustness

Mean LIs reported from sentence generation studies are illustrated in Figure 3.2. High mean LIs of between 0.74 and 0.89 have been reported for sentence generation, both when sentences are pre-learnt prior to scanning (e.g. Stippich et al., 2003), and when they are actively generated during the scan (e.g. Tzourio-Mazoyer, Joliot, Marie, & Mazoyer, 2016). However, other studies have reported more modest laterality estimates of between 0.48 and 0.65, again with both variants of the task (Mazoyer et al., 2014; Partovi et al., 2012a; Partovi et al., 2012b). Thus, it does not appear to be the case that strength of laterality differs according to whether sentences are generated spontaneously during the scanning session, or learnt prior to scanning. Two studies within the current literature search measured laterality for semantic fluency and sentence generation within the same participants (Partovi et al., 2012a; Stippich et al., 2003); however, these studies reported differences in strength of laterality between the tasks in different directions. Further, they used the version of the sentence generation task in which sentences are learnt prior to scanning, meaning task demands were not well matched. Thus, there is currently insufficient data with these tasks to evaluate predictions of stronger laterality for sentence over word processing. Partovi et al., (2012a) report good reproducibility for sentence generation, however their analysis simply looked at significant differences in group means over repeated testing, and not at reproducibility of individual participant's LIs.

3.3.2.2 Effect of ROI

There is mixed evidence as to whether sentence generation yields differences in the laterality measured from frontal versus temporoparietal ROIs. While Partovi

et al. (2012a; 2012b) and Stippich et al., (2003) reported equivalent strength of laterality across both, Tzourio-Mazoyer et al. (2016) reported significantly stronger laterality in frontal than temporal areas. Interestingly, in contrast to the exclusively right handed samples of the former studies, this latter study used a mixed handedness sample, with an overrepresentation of left handers. This suggests the possibility that greater regional heterogeneity may characterise the atypical profiles of language lateralisation that are more often found within atypical handedness samples.

3.3.2.3 Effect of baseline

When sentences are learnt prior to scanning, studies generally employ simple cross-fixation or rest as a baseline (Partovi et al., 2012a; Partovi et al., 2012b; Stippich et al., 2003). Conversely, the two studies using spontaneous generation of sentences during the scanning session used an active linguistic baseline, in which participants covertly generated the months of the year (Mazoyer et al., 2014; Tzourio-Mazoyer et al., 2016). A comparison of these baselines in terms of the language processes isolated by each contrast is given in Table 3.3. As can be seen, the active baseline subtracts out activity related to speech motor planning to leave those processes specific to the construction of novel sentences, such as syntactic and lexico-semantic processing; conversely, the contrast with rest results in poor isolation of such language processes. This highlights the need to consider carefully the functions one wishes to isolate in choosing a suitable baseline, and the implications this will have for interpretation of measured laterality in relation to linguistic processes.

Table 3.3: Language processes isolated by subtraction of different baseline tasks from sentence generation. *Table shows comparison of passive and active baseline tasks as used for subtraction with sentence generation, in terms of the language and domain general processes engaged by each paradigm and isolated by the subtraction contrast. ✓= engaged,* ***(✓)*** *= sometimes engaged (e.g. depending on task demands, modality of stimuli, occurrence of automatic linguistic processing).*

Contrast	Speech motor planning/ articulation	Phono-logical access	Ortho-graphical processing	Semantics	Word retrieval	Syntax	Working memory	Motor processing	Auditory processing
Task: Sentence generation	✓	✓		✓	✓	✓	✓		
Baseline: Rest				**(✓)**					
Sentence generation vs Rest	✓	✓		(✓)	✓	✓	✓		
Task: Sentence generation	✓	✓		✓	✓	✓	✓		
Baseline: Recite months	✓	✓		**(✓)**	**(✓)**				
Sentence generation vs recite months				(✓)	(✓)	✓	✓		

3.3.3 ***Passive speech listening***

Passive speech listening as a language paradigm appears to yield more variable laterality estimates, perhaps reflective of the wide variety of language processes that it can engage (see Table 3.2). Lower level acoustic processing of speech sensory input is predicted to be bilateral by multiple models (Hickok & Poeppel, 2007; Peelle, 2012; Poeppel, 2014; Price, 2012). However, there are discrepancies in the extent of lateralisation assumed for mapping of sound to meaning, considered bilateral by Hickok and Poeppel (2007) but left lateralised by other authors (Peelle, 2012; Poeppel, 2014; Price, 2012) due to the need to process meaning at a sentential level. Thus, depending on the baseline subtraction used, different levels of processing may be isolated to result in variable levels of laterality.

3.3.3.1 LI strength, reliability and robustness

The majority of studies using passive listening tasks report very weak average LIs (see Fig. 3.3 for mean LI values), indicating bilateral activation, and it is often the most weakly lateralising task when compared to others (Binder, Swanson, Hammeke, & Sabsevitz, 2008; Harrington et al., 2006; Miro et al., 2014; Ocklenburg et al., 2013; Tzourio-Mazoyer et al., 2015). A notable exception to this was presented by Thivard et al. (2005) who reported a mean laterality index of 0.72 within a frontal ROI for a passive story listening task, stronger than that seen in this ROI for a semantic fluency task (0.51). High test-retest correlations have been reported for speech listening (Razafimandimby et al., 2007).

3.3.3.2 Effect of ROI

Studies are inconsistent as to whether stronger laterality is found for speech listening within a frontal or a temporoparietal ROI. Harrington et al. (2006) reported that temporoparietal LIs were stronger and more reliable than frontal LIs, whereas other studies have reported weaker and more variable LIs for a temporal compared to a frontal ROI (Miro et al., 2014; Ocklenburg et al., 2013; Thivard et al., 2005). Thus, despite being a receptive task, in general posterior language areas appear to be poorly lateralised for speech listening tasks (although this may depend on the baseline task employed - see *3.3.3.3 Effect of baseline task*).

3.3.3.3 Effect of baseline task

The varying levels of asymmetry reported for speech listening tasks may in part be attributable to the baseline used by different studies. Binder et al. (2008) found that changing the baseline from rest to tone listening raised the average LI for word listening from 0.1 to 0.52. In this regard it is interesting that the two studies reporting near-zero average LI values for speech listening employed rest as a

baseline task (Miro et al., 2014; Ocklenburg et al., 2013). Conversely, Thivard et al. (2005) and Harrington et al. (2006) both used backwards speech listening as a baseline and reported stronger laterality measurements for speech listening. This effect of baseline is consistent with the idea of bilateral early auditory processing that must be subtracted out by a non-linguistic auditory stimulus in order to reveal asymmetry for higher-level 'central language processes' (Peelle, 2012; Poeppel, 2014; Price, 2012).

3.3.4 ***Text reading***

Reading text or narrative requires decoding of orthography into phonological representations, semantic and syntactic processing of the decoded sentence, and binding within and across sentences to arrive at an overall understanding of text meaning (see Table 3.2). Visual word form processing is considered to rely on a lateralised ventral occipito-temporal region, although this may not reflect a left specialisation for orthography *per se* (Price, 2012).

3.3.4.1 LI strength and effect of ROI

The current literature search identified only two papers investigating lateralisation of text reading. Both studies used the same covert (silent) text reading task with a baseline of covert reading of text composed of pronounceable non-words. Backes et al. (2005) reported moderately strong laterality (LI = 0.59) using a combined frontal-temporoparietal ROI, whereas (Deblaere et al., 2002) reported weak laterality (LI = 0.21) using a global LI (see Fig. 3.3). This supports the hypothesis that global LIs are generally weaker than regional LIs.

3.3.5 ***Phonemic judgement***

Phonemic judgement tasks require a decision relating to phonological structure; most commonly, a rhyme judgement. This task relies on mapping of acoustic or visual input onto phonological units such as phonemes and syllables, a process known as decoding. The precise nature of these stored phonological codes remains a debate; according to theorists in the tradition of the motor theory of speech perception (e.g. Liberman & Mattingly, 1985), these are represented as speech motor gestures in left premotor cortex. Hickok and Poeppel (2007) argue that while the phonological codes themselves are bilaterally represented, the process of their mapping onto articulatory motor representations relies on a left lateralised dorsal stream. Conversely, other models propose that such processing of single words is a less strongly lateralised process (Peelle, 2012).

3.3.5.1 LI strength, reliability and robustness

Phonemic judgement tasks yield relatively strong laterality, with reported LI values ranging from 0.41 to 0.84 (see Fig. 3.2). However, when compared to other tasks, phonemic judgement is often reported as more weakly lateralising (Baciu, Watson, et al., 2005; Niskanen et al., 2012; Seghier et al., 2004). Phonemic judgement may be superior to other tasks however in terms of robustness and reproducibility. Morrison et al. (2016) reported that a rhyming decision task demonstrated greater reliability than a word generation task, yielding reproducible dominance classifications in 100% of participants, and average test-retest correlations for LI values of 0.9 and above. Furthermore, such reproducibility of LIs obtained with rhyming decision was more robust against changes in the activity measure used for LI calculation.

3.3.5.2 Effect of ROI

Rhyming decision tasks yield particularly strong laterality when a frontal or a combined frontal-temporoparietal ROI is used (Baciu, Watson, et al., 2005; Clements et al., 2006; Cousin et al., 2007; Niskanen et al., 2012). For example, Cousin et al. (2007) identified a particularly strong leftward asymmetry for the inferior frontal gyrus during rhyme detection. Thus, frontal ROIs may be optimal for yielding the strongest laterality with this task.

3.3.5.3 Effect of baseline task

All studies within the current literature search using phonemic judgement were found to employ an active perceptual decision baseline task on either non-linguistic material (e.g. line orientation matching) or nonsense words or characters (e.g. nonsense word font matching). This subtracts out non-linguistic working memory processes (see Table 3.2), as well as basic visual processing.

3.3.6 ***Semantic decision***

Semantic decision tasks require a judgement about a word's semantic content or about the semantic relationship between a pair of words, such as whether two words belong to the same category. Such conceptual knowledge is proposed to rely on a distributed processing network, with different brain areas each contributing to different aspects of an item's representation (Warrington & McCarthy, 1983, 1987; Warrington & Shallice, 1984). In addition to this distributed network, Patterson, Nestor, and Rogers (2007) have argued for the existence of a 'semantic hub' within the bilateral anterior temporal lobes that integrates the distributed modality-specific representations into one amodal representation. However, a recent meta-analysis of functional imaging studies by Rice, Ralph and Hoffman (2015) suggested that while conceptual knowledge does appear to

be represented bilaterally in the anterior temporal lobes, left lateralised activity was more likely when semantic content was accessed linguistically. This is in contrast to the predictions of Hickok & Poeppel's (2007) model of language in which access to lexico-semantics from speech processing (via the ventral stream) is considered as a bilateral process.

3.3.6.1 LI strength, reproducibility and robustness

The strength of laterality reported for semantic decision tasks is quite variable, ranging from near-zero to around 0.8 (see Fig. 3.3 for mean LI values). This may depend on the type of semantic decision required by the task. Tasks which require judgement of the semantic relatedness of two words appear to yield relatively strong laterality, ranging from 0.59 to 0.84 (Bethmann et al., 2007; Fernandez et al., 2001; Häberling et al., 2016; Seghier et al., 2004). In contrast, category membership tasks with single words appear to give much lower LIs, ranging from 0.03 to 0.52 (Deblaere et al., 2002; Hund-Georgiadis et al., 2002; Hund-Georgiadis, Lex, & von Cramon, 2001; Ramsey et al., 2001; van Oers et al., 2010). This suggests that it may be the process of integrating and comparing across semantic representations for different concepts that is strongly lateralising; conversely a simple lexical look-up to determine the category membership of a single concept may not be strongly lateralised.

Jansen et al. (2006) reported very low reproducibility for synonym decision LIs across a range of different LI calculation methods within Broca's area; much higher reproducibility was found however when a temporoparietal ROI was used. Conversely, Harrington et al. (2006) reported high test-retest correlations for an abstract/concrete semantic decision task across both frontal (IFG) and

temporoparietal ROIs. This discrepancy between the two studies could be due to the differences in the tasks they used.

3.3.6.2 Effect of ROI

The majority of studies report no significant differences in the magnitude of laterality found within temporoparietal and frontal ROIs for semantic decision tasks (Bethmann et al., 2007; Häberling et al., 2016; Harrington et al., 2006; Hund-Georgiadis et al., 2002; Ramsey et al., 2001; van Oers et al., 2010). Some studies have reported differences across ROIs, however these can be in opposite directions (Fernandez et al., 2001; Szaflarski et al., 2008). As discussed, some evidence suggests that LIs calculated from temporoparietal ROIs for semantic decision may be more reproducible than those calculated from frontal ROIs (Jansen et al., 2006).

3.3.6.3 Effect of baseline tasks

Semantic decision laterality is also strongly influenced by the baseline task used. Binder et al. (2008) and Hund-Georgiadis et al. (2002, 2001) manipulated the baseline and found that the use of an active perceptual decision task as opposed to passive rest yielded a large increase in the strength and consistency of semantic decision laterality. Binder et al. (2008) argued that resting baselines are unsuitable for subtraction with semantic decision, since they allow for the activation of conceptual language representations as the participant ‘day dreams’ and engages in ‘inner speech’. An active perceptual decision baseline interrupts such ongoing conceptual processing and engages the same executive and attentional processes as the language paradigm. This subtraction is illustrated in Table 3.4, which illustrates the better isolation of semantic processes that this baseline provides compared to the contrast with rest. Baseline tasks that engage

linguistic processing themselves may result in reduced laterality; for example, Deblaere et al. (2002) suggested their finding of weak laterality for semantic decision may have been due to a vowel decision baseline task. However, it should be noted that Binder et al. (2008) reported identical laterality strength (a mean LI of 0.62) for semantic decision using either a baseline of tone decision or phoneme decision. Overall, this evidence suggests that baseline tasks used for semantic decision must be active, sufficiently engaging and challenging so as to prevent 'day-dreaming', and ideally involve material from a non-linguistic domain e.g. symbols or tones.

Table 3.4: Language processes isolated by subtraction of different baseline tasks from semantic decision. *Table shows comparison of active and passive baseline tasks as used for subtraction with semantic decision, in terms of the language and domain-general processes engaged by each paradigm and isolated by the subtraction contrast.* ✓*= engaged,* **(✓)** *= sometimes engaged (e.g. depending on task demands, modality of stimuli, occurrence of automatic linguistic processing).*

Contrast	Speech motor planning/ articulation	Phono-logical access	Ortho-graphical processing	Semantics	Word retrieval	Syntax	Working memory	Motor processing	Auditory processing
Task: Semantic decision		**(✓)**	**(✓)**	✓	**(✓)**		✓	**(✓)**	
Baseline: Rest				**(✓)**					
Semantic decision vs Rest		(✓)	(✓)	**(✓)**	(✓)		✓	(✓)	
Task: Semantic decision		**(✓)**	**(✓)**	✓	**(✓)**		✓	**(✓)**	
Baseline: Perceptual decision							✓	**(✓)**	
Semantic vs Perceptual decision		(✓)	(✓)	✓	(✓)				

3.3.7 ***Sentence comprehension***

Sentence comprehension tasks require some judgement about the content of a spoken or written sentence. Syntactic and semantic processing are often confounded (see Table 3.2); for example, the task may require participants to

decide if two sentences with different grammatical constructions have the same meaning. However, they are noteworthy among other tasks in the extent of their syntactic processing requirements. Laterality of syntax has been a subject of debate, with some authors arguing for a bilateral involvement in syntax (e.g. Hund-Georgiadis et al., 2001), but others arguing for a left dominance (Friederici, 2011; Tyler et al., 2011; Wright, Stamatakis, & Tyler, 2012). At a more general level, multiple models would predict left lateralisation for the sentence-level processing engaged by this task, without making specific claims about lateralisation of syntactic processing *per se* (e.g. Peelle, 2012; Poeppel, 2014; Price, 2012).

3.3.7.1 LI strength, reliability and robustness

Multiple studies report strong laterality for sentence comprehension tasks (Harrington et al., 2006; Jensen-Kondering et al., 2012; Kennan, Kim, Maki, Koizumi, & Constable, 2002; Niskanen et al., 2012; Sanjuan, Forn, et al., 2010b; Vassal et al., 2016), with LI values ranging from 0.55 to 0.88 (see Fig. 3.2). Studies which compare sentence comprehension laterality measures to those of other tasks suggest that it can outperform semantic decision, phoneme decision, story listening and naming tasks in terms of the strength of laterality, although this can depend on the ROI (Harrington et al., 2006; Niskanen et al., 2012).

3.3.7.2 Effect of ROI

Evidence appears inconsistent as to the effect of ROI on the laterality obtained with sentence comprehension tasks. Studies have reported both stronger laterality for frontal than temporoparietal ROIs (Jensen-Kondering et al., 2012; Niskanen et al., 2012) and vice versa (Harrington et al., 2006; Sanjuan, Forn, et al., 2010b). In terms of reliability of lateralisation, Harrington et al (2006) reported

very high reproducibility for a visual sentence comprehension task across both frontal and temporoparietal ROIs, with test-retest correlations above 0.9. In contrast, auditory sentence comprehension yielded more reliable lateralisation within a temporoparietal than a frontal ROI. Modality of the stimuli may thus affect which ROI is optimal.

3.3.7.3 Effect of baseline task

Generally, active baselines are employed for sentence comprehension tasks. A notable exception is seen in Harrington et al. (2006) who used passive listening to backwards speech as a baseline for auditory sentence comprehension. This passive baseline may explain the weaker laterality they reported as compared to other studies (mean LI of around 0.45). Interestingly, Sanjuan et al. (2010) who also reported a relatively low level of laterality compared to other studies used phoneme decision as a baseline. As previously discussed in relation to a study by Deblaere et al. (2002) using semantic decision, it is possible that the use of such a baseline with high linguistic processing demands may lower the strength of the laterality seen.

3.3.8 ***Naming***

Naming tasks require the generation of the name of an item in response to either a visual (pictorial) or verbal description. According to Hillis' (2007) model of naming, picture naming involves three major levels of processing; a semantic level in which amodal general and specific semantic information is accessed from a structural description of an object; a lemma level which involves the defining features of an object at a more abstract level (e.g. what makes a sheep a sheep); and a phonological/orthographical level, in which the phonological and orthographical representations associated with that concept are accessed. Thus,

naming tasks have the potential to engage multiple key components of the language network (see Table 3.2).

3.3.8.1 LI strength, reliability and robustness

Studies using naming as a language activation task report a wide variety of LI values, ranging from 0.08 to 0.96 (see Fig. 3.2). This can partly be explained by variation in the nature of the naming task used. Zero to moderate laterality has been reported for picture naming tasks, which are often the least lateralising when compared to other tasks (Deblaere et al., 2002; Harrington et al., 2006; Jansen et al., 2006; van Oers et al., 2010; Vikingstad et al., 2000). However, the lateralising ability of naming tasks can be increased by the addition of sentence comprehension demands. A naming from written or auditory description task known as 'responsive naming' requires comprehension of a question describing an object in order to generate the required name. This has been reported to yield strong laterality across both frontal and temporal ROIs, in the range of 0.65 to 0.96 (Gaillard et al., 2002; Niskanen et al., 2012). This increase in laterality with the addition of sentence-level processing is consistent with models of language which predict an increase in laterality for connected versus unconnected language i.e. structured sentences versus single words (Peelle, 2012; Poeppel, 2014).

Picture naming also shows poor reliability in laterality measurement. Jansen et al. (2006) reported that picture naming did not determine dominance reproducibly in about a third of participants. Rutten et al. (2002) similarly reported a failure to find significant test-retest correlations for naming LIs. Significant test-retest correlations were reported by Harrington et al. (2006) for a picture naming task at around the same level as those seen for semantic decision; however,

reproducibility of naming laterality was lower than that seen for verb generation and sentence comprehension.

3.3.8.2 Effect of ROI

Naming tasks do not appear to favour one ROI over another in laterality measurement. The current literature search found two studies which reported differences in the laterality measured from frontal and temporoparietal ROIs, however this difference was in opposite directions (Harrington et al., 2006; Brennan et al., 2007). The majority of studies instead reported highly similar strength of laterality across frontal and temporoparietal ROIs (Gaillard et al., 2002; Niskanen et al., 2012; van Oers et al., 2010; Vikingstad et al., 2000). Furthermore, Rutten et al. (2002) reported similar levels of reproducibility of naming LIs for both regions.

3.3.8.3 Effect of baseline task

The baselines used with picture naming provide interesting evidence on the processes underlying its laterality. Deblaere et al. (2002) reported near zero laterality for picture naming using a baseline which required participants to name the position of the intersection of four lines (e.g. up, down, left, right). This task involves engagement in similar semantic and word retrieval processes (see Table 3.2), which may explain the weak laterality measured. However, Brennan et al. (2007) reported strong laterality for a picture naming task with a number counting baseline. This would subtract out speech production and word retrieval processes for an automated speech sequence, predicted to involve bilateral activity (Price, 2012; Poeppel, 2014). This subtraction would thus isolate spontaneous non-automated retrieval and word generation processes which may engage left hemisphere language systems, increasing measured laterality.

3.3.9 ***Combined Task Analysis***

Combined task analysis (CTA) involves the calculation of LIs from contrast images generated by combining scans across multiple language tasks. This method identifies commonalities between tasks' activity patterns in order to isolate the 'core' language network, and exclude task-specific, non-linguistic activity caused by differences in task design that may influence the LI value. In this way, CTA can represent a theoretical alternative to baseline tasks to subtract domain-general activity, assuming that different tasks involve different patterns of non-linguistically relevant activity. Indeed, there is evidence that CTA results in higher and more reliable and robust estimates of laterality for language (Dodoo-Schittko et al., 2012; Harrington et al., 2006; Jansen et al., 2006; Niskanen et al., 2012; Ramsey et al., 2001; Rutten et al., 2002; Sommer, Ramsey, Mandl, & Kahn, 2003; van Rijn et al., 2008).

However, the theoretical assumptions which motivate CTA can be questioned. CTA assumes that variability in laterality for different tasks should be ascribed to non-linguistic processes or viewed simply as measurement error, rather than reflecting the underlying nature of hemispheric organisation for language (e.g. Ramsey et al., 2001). Such a theoretical stance ignores the possibility of multidimensional lateralisation across different language processes. Indeed, recent fMRI studies have reported cases of dissociated laterality for different language functions within individuals (see *Chapter 2: 2.3.6.3 Regional heterogeneity in laterality* for a review), corroborating those early clinical reports of a 'division of labour' across the hemispheres in patients with bilateral language representation (Rasmussen & Milner, 1975).

Such differences in dominance between language tasks would be lost in a CTA, since combining scans across paradigms would result in few areas of common activation and thus a loss of these tasks differences. Instead, it will be necessary to design fMRI protocols that probe the within-subject variation in language lateralisation across a range of tasks, while controlling for non-linguistic confounds. Conversely, the strong, reliable and robust LI values provided by CTA would be more useful in cases where a clear categorical decision on an individual's language lateralisation is required.

3.4 Summary and conclusions

The systematic review presented in this chapter has highlighted the high level of variation and inconsistency in the strength and reliability of laterality measured using different language tasks. As per hypotheses, some of this variability in laterality is related to parameters such as the region of interest and baseline task used, which can have task-specific effects. In general however, the current state of the literature is such that it is difficult to draw clear conclusions that can be used to guide task selection. This review highlights the need for more research that systematically compares laterality across different tasks in within-subject designs, with rigorous matching of non-linguistic aspects of task design.

This review of the literature does suggest however some practical recommendations that can be used to guide task design. Extensive use of verbal fluency is clearly warranted given the robustness of its lateralisation; however, the common employment of passive baselines should be replaced with active baselines carefully chosen so as to isolate the language process of interest whilst controlling for all other processes. Comparison of word generation with sentence generation offers the opportunity to test predictions of models that assume

stronger laterality for processing of connected sentences over single words (Peelle, 2012; Poeppel, 2014). Future studies should try to more closely match task demands of word generation and sentence generation tasks, in order to systematically compare their strength of laterality within subjects. For example, ideally one would use the same stimuli for each task such as names of different categories, with a cue indicating whether the subject should generate instances of these categories (semantic fluency) or generate a sentence taking an instance of this category as its head noun.

Evidence on semantic decision paradigms suggests that stronger laterality can be observed when tasks require integration across the semantic content of different concepts (e.g. semantic relatedness decision), rather than simple category membership decision on single words. Similarly, where naming tasks are used, evidence suggests that naming from description yields more robust laterality measurement than naming from pictures; that is, an additional sentence comprehension component appears to improve the lateralising power of this task. Sentence comprehension tasks themselves appear to yield strong laterality; however, more work is needed to develop such tasks in order to attempt to disentangle semantic and syntactic components. Indeed, this review highlights a distinct lack of tasks in language laterality research that aim to primarily engage syntactic processing, reflected in the lack of consensus over the strength of its hemispheric specialisation. Further work is needed to design and validate tasks that isolate syntactic processing for laterality measurement. One possibility might be offered by tasks involving judgements on ‘jabberwocky’ sentences (e.g. Fedorenko, Nieto-Castanon, & Kanwisher, 2012) in which content words are

replaced by non-words (thus preserving syntactic structure but removing semantic content).

More work is needed to investigate the potential significance of variability in laterality across different language functions, both within individuals and at a group level. Growing appreciation of the potential significance of cases of dissociated dominance, both in clinical and healthy samples, should encourage the field to move away from the use of single tasks and single ROIs. Instead, research should focus on developing batteries of closely matched tasks that tap a variety of language functions to allow systematic comparisons in within-subject studies. This will ultimately allow for more quantitative meta-analyses of such literature, to draw stronger conclusions as to patterns of laterality across different components of the language network.

One way to approach this would be to develop a generic task format in which the participant is always performing the same form of task with the same type of stimuli but with regards to different linguistic parameters. For example, one such format could be a decision task in which one must decide if pairs of word stimuli are ‘matching’ or ‘non-matching’. The parameters that define matching and non-matching pairs can then be varied according to the language process of interest; for example, rhyming versus non-rhyming (phonology), same semantic category or different semantic category (semantics) or same syntactic category or different syntactic category (syntax). These could be interleaved with another visual cue indicating which decision should be made on the current trial. In this way, a generic task format would remove non-linguistic differences in task design that can confound interpretation of differences in laterality. This type of approach has been more fully developed by Price and colleagues in the development of a

battery of tasks devised for a within-subject fully balanced factorial design, with tasks corresponding to all possible combinations of levels of factors relating to experimental design aspects (e.g. stimulus modality, linguistic content, form of response). This has been used to test contrasts that allow fractionation of different levels of linguistic processing for localization of brain activity (e.g. Hope et al., 2014); future work could implement a similar battery of balanced tasks for lateralisation measurement. The current thesis attempts to take this approach to measurement of laterality in a study presented in Chapter 5, where matched phonological decision and semantic decision tasks are employed.

CTA can provide an efficient method of isolating language activity shared across multiple different aspects of language functioning, to allow robust and reliable measurement of laterality of the core language network. However, this methodology appears to be motivated by an implicit assumption pervasive across much laterality research that there is a single core language network which displays a unitary and perfect lateralisation; thus the ability of an fMRI protocol to provide a good measure of language laterality depends on its ability to uniquely engage this language network and to yield a laterality index of close to 1. Tasks which yield LIs further from 1 therefore are viewed as inadequate measures of language lateralisation.

However, defining the sensitivity of a task to capture the ‘true’ lateralisation of a language function in terms of the strength of its laterality can be challenged. Such an approach would lead one to reject tasks that yield lower LIs, which may in fact reflect meaningful variation in hemispheric organisation within the language network. For example, naming is a complex linguistic function that requires both receptive and expressive components, with access to both semantics and

phonology; however the evidence reviewed here shows it yields low LI values (e.g. Deblaere et al., 2002). Is it appropriate to conclude that naming is therefore an 'impure' measure of language network function? Or rather, could this tell us something about the hemispheric organisation of the language functions on which it relies? I would argue that research should be open to the possibility that it may be possible to validly and reliably measure laterality for a language process, and yet still obtain a low LI.

This raises the question of how one should judge laterality paradigms; what metric should one use for judging 'success' in accurately measuring an individual's laterality? This chapter has highlighted how different methods of laterality measurement can result in variable LI values for an individual across different regions, active language tasks and baseline tasks. For example, in the case of verbal fluency tasks, an individual may show a stronger LI when an active rather than a passive baseline is used (Dodoo-Schittko et al., 2012). How should one then decide which of these can be considered to best reflect the 'true' laterality of an individual? In this case, the greater strength of laterality with an active baseline is often taken to indicate that this is a more accurate laterality measurement; however, other metrics such as the reliability of the laterality and in clinical work its predictiveness of post-surgical outcomes may arguably represent better standards for assessing goodness of laterality measurement. In this way, it will be important for the field to consider more deeply the metrics that are used to compare the relative utility of LIs yielded by different paradigms, and to challenge the implicit 'strongest is best' assumption that commonly guides interpretation of task LI values.

Chapter 4 Profile of language abilities in a sample of high functioning adults with developmental disorders

Abstract

This study aimed to investigate the profile of abilities in a sample of high functioning English-speaking adults with a history of developmental disorders. A total of 104 adult participants were recruited, 67 of whom had a diagnosis of a developmental disorder (including dyslexia, dyspraxia and autism) and 37 of whom had no diagnosis. Participants with a developmental disorder were found to demonstrate impairments across measures of language, literacy and working memory. When considering specifically those participants with a diagnosis of dyslexia (51 participants) compared to those without (53 participants), impairments were found on measures of word reading, phonological skills, processing speed, working memory and oral language skills. Conversely, spared functioning was seen for non-verbal intelligence, vocabulary, story comprehension and sentence construction. Logistic regression analyses found that a measure of phonological skills alone was optimal for discriminating between these groups. An exploratory data-driven clustering analysis found that diagnosis was not a reliable predictor of similarity between cases, with diagnostic categories split between data-driven clusters. This highlights the importance of assessing support needs on a case-by-case basis, rather than according to diagnostic label. Overall, despite being a 'high functioning' sample, a large

number of participants within this sample still demonstrated impairments on skills that are likely to affect success in higher education.

4.1 Introduction

Despite the large amount of research that has gone into characterising the profiles of impairment in children with diagnoses of developmental disorders, little has been done to follow these children up at adulthood. More and more adults with developmental disorders are entering higher education; in particular, there have been large increases in the number of dyslexic students entering University (P. Henderson, 2017; Mortimore & Crozier, 2006). This may be because of better identification and support at school, or better facilities available at institutes of higher education such as disability support services. Further, it may be that these individuals, although encountering difficulties earlier in education, have sufficiently compensated or 'recovered' from these problems. It has been reported that such compensation to achieve reading skills in the normal range by adulthood occurs in around 25% of individuals who experience reading difficulties as a child (Lefly & Pennington, 1991). This raises the question as to the level of support such individuals should receive during their higher education. That is, do adults who received a diagnosis of a developmental disorder as a child still experience difficulty within their affected areas?

4.1.1 *Dyslexia in adulthood and higher education*

There are multiple motivations for characterising so called 'compensated' or 'highly functioning' adult dyslexics who have successfully entered higher education. Firstly, identifying what enables some individuals with a history of reading difficulties to compensate to a point sufficient to succeed at university will provide useful insights for what areas of support should be targeted in the school

years. Secondly, understanding whether these 'compensated' individuals nevertheless still experience persisting deficits in areas of known weakness in childhood is important for considering whether and how these students should be supported during their studies, to ensure they are not at a disadvantage for academic success. A recent study by Bergey, Parrila and Deacon (2018) reported that students entering a Canadian university with a history of reading difficulties demonstrated lower academic self-efficacy and achieved lower grades than their typically reading peers. This suggests that these highly compensated individuals may still be at a disadvantage when entering higher education.

There have been a number of studies aiming to characterise the profile of impairment demonstrated by adults with a history of reading disorders. A seminal meta-analysis of this literature by Swanson and Hsieh (2009) found that adults with reading disabilities demonstrated a range of deficits on standardised measures that extended beyond a phonological core deficit, for example with additional impairments on naming speed, verbal memory, writing and maths skills. Age was found to be unrelated to the magnitude of effect sizes across studies, suggesting that deficits in adults with reading disorders are persistent across age.

Other studies, however, have aimed to take a more positive approach by considering what skills may be preserved or even superior in highly compensated dyslexic university students, which may have enabled their success in entering higher education. A sparing of vocabulary skills is often reported (e.g. Hatcher, Snowling, & Griffiths, 2002), with one study even reporting superior performance on a measure of vocabulary depth in French speaking dyslexic students compared to controls (Cavalli et al., 2016). A strong vocabulary may scaffold

better reading comprehension in the face of poor decoding skills. A study by Cavalli, Duncan, Elbro, El Ahmadi and Cole (2017) reported a dissociation between impaired phonological and intact morphological skills in a dyslexic student sample; further, the higher this discrepancy within an individual, the better their reading. It was suggested that intact morphological awareness may provide adult dyslexics with an alternative reading strategy; the success of this however may be heightened within this French speaking sample, since the French writing system is largely morphologically based. Overall, these more semantic based skills may offer compensatory strategies for reading that can enable dyslexics to access and cope with the large amounts of written information they will encounter during a university degree.

Despite these reports of areas of relative strength in adult dyslexics, a recent study suggested that dyslexic students may even demonstrate impairments that extend to spoken language. Wiseheart and Altmann (2018) demonstrated that high functioning dyslexic university students were impaired on a spoken sentence generation task relative to controls, being slower to respond and constructing sentences that were less precise and less fluent. This impairment was exacerbated by poorer working memory, but mitigated by strong vocabulary. This demonstration of a deficit in spoken language abilities in compensated dyslexics is striking, and suggests that a range of difficulties not limited to the traditional core phonological deficit still persist in adult dyslexics who have successfully reached higher education.

4.1.2 *Profile of abilities in adults with developmental disorders*

The current study aimed to further investigate the profile of abilities and impairments in a sample of high functioning adults with a prior diagnosis of a

developmental disorder, with a particular focus on dyslexic adults. This was firstly to provide a detailed characterisation of a sample of adults with a range of developmental disorders (*Aim 1*) which was the subject of study in a further chapter of this thesis (*Chapter 5*). In that study, I tested for the presence of a relationship between individual variability in lateralisation patterns and language abilities; the analyses presented here thus aimed to demonstrate that this sample did contain a wide range of language abilities, with the prediction that individuals with developmental disorders would be significantly impaired across the measures.

Secondly, this chapter seeks to determine which measures were particularly sensitive for discriminating the current high functioning dyslexic group from the rest of the sample, including both controls and other developmental disorders (*Aim 2*). The rationale for comparing dyslexic participants to 'non-dyslexic' participants both with and without other developmental disorders was to try to identify what areas of weakness specifically characterise those participants with a diagnosis of dyslexia in the current high functioning sample, as opposed to being common across different diagnoses. Whilst the unusual nature of this sample means that the findings of this analysis are of very limited generalisability to the wider population with dyslexia, the aim was to provide further description of the nature of the difficulties experienced by our 'compensated' dyslexic participants, who constituted a large portion of the developmental disorder group. Based on the above reviewed literature, I predicted that impairments would not only be demonstrated on measures directly relating to phonological skills, but further extend to other measures such as processing speed, working memory and oral language. In contrast, based on this previous evidence I predicted this

sample would demonstrate spared abilities in the domains of vocabulary and non-verbal IQ (e.g. Hatcher et al., 2002).

Thirdly, this chapter aimed to explore learning profiles in this heterogeneous sample in a more data-driven manner (*Aim 3*). Research on developmental disorders typically focuses on single diagnostic groups, with strict exclusionary criteria for recruitment of 'pure' cases of a specific disorder. Increasing awareness of the need to consider heterogeneity within and comorbidity across diagnostic groups has led to a small number of recent studies that have explored alternative data-driven methods for deriving cognitive and behavioural profiles in developmental samples (Archibald, Cardy, Ansari, Olino, & Joanisse, 2019; Archibald, Cardy, Joanisse, & Ansari, 2013; Astle, Bathelt, & Holmes, 2018; Bathelt, Holmes, et al., 2018; see Hong, 2018 for an editorial comment). These studies have applied unsupervised machine learning approaches to data obtained from heterogeneous samples containing a range of difficulties and diagnostic labels. The key feature of this approach is that the algorithm is not fed pre-specified group memberships but instead attempts to derive sub-groups based on similarity structure within the data. This research has highlighted that such data-driven 'clusters' do not necessarily correspond well to diagnostic labels of cases. In order to further explore the profiles of impairment present in my highly heterogeneous sample of adults, this chapter aimed to apply such an unsupervised machine learning approach to data obtained from a sub-set of measures most sensitive to discriminating dyslexic from non-dyslexic participants.

4.1.3 *Dyspraxia in adulthood*

One question of interest for this data-driven clustering analysis relates to whether cases of dyspraxia would tend to cluster with cases of dyslexia or with controls when considering those measures most relevant to dyslexia. Dyspraxia[1] or DCD (developmental coordination disorder) is a developmental disorder primarily affecting motor coordination; however, the term is often used quite loosely to further incorporate problems in executive functioning and in some cases language and reading. It is known to frequently co-occur with other disorders such as ADHD and dyslexia (Kadesjö & Gillberg, 1999). Compared to these disorders however, dyspraxia in childhood has received much less research focus, with even fewer studies of adults. A study by Dewey, Kaplan, Crawford and Wilson (2002) reported that children meeting criteria for DCD demonstrated a range of difficulties across measures of attention, reading, writing and spelling; in particular, this study highlighted a particular difficulty with non-word reading in this sample, indicative of phonological processing difficulties. The overlap between dyslexia and dyspraxia as suggested by such findings has likely led to cases of mislabelling of children being referred for learning difficulties. A professional consensus on recognising DCD as a developmental disorder was only reached relatively recently (Polatajko, Fox, & Missiuna, 1995), and a lack of agreement on the primary characteristics (e.g. purely motoric versus more general issues with organisation) continues (Sugden, 2006). Kirby, Sugden, Beveridge, Edwards and Edwards (2008) argued that 15 to 20 years ago when

[1] For clarification, the term dyspraxia here is not used to refer to verbal dyspraxia, but the general motor coordination disorder also known as DCD.

the disorder dyspraxia was less well known, a child presenting with such symptoms would have likely been given the label of dyslexia.

This overlap in symptomology and likely misapplication of diagnostic labels raises the question of whether students entering higher education with these different diagnoses will face similar difficulties and need similar types of support to succeed at university. Kirby et al. (2008) compared groups of students in higher education with either a diagnosis of dyspraxia, dyslexia or comorbid dyspraxia and dyslexia on various self-report measures. Students with dyspraxia were more likely to report executive functioning difficulties (e.g. poor organisation, high distractibility) and to live at home with parents than students with dyslexia. Students with dyslexia however were more likely to be in receipt of Disabled Students' Allowance (DSA) than students with dyspraxia. Interestingly there were no differences between diagnosis groups in the type of support received (e.g. extra time, help with organisation), suggesting a 'one size fits all' approach to learning disabled student support. It is unclear however to what extent the profile of impairment presented by high functioning dyspraxic versus dyslexic adults overlaps; for example, it is possible that despite similar profiles in childhood, the patterns of strengths and weaknesses shown by individuals with these disorders may diverge in adulthood.

A lack of comparative studies of these two disorders in adult samples has meant a gap in knowledge as to what types of difficulties are likely to persist in dyspraxic adults and how similar these are to profiles of impairment in adult dyslexics. The present study aimed to provide preliminary insight into such questions (albeit with a small and unusual sample) by considering the extent to which adults with a diagnosis of dyspraxia cluster together with adults with a diagnosis of dyslexia

using measures relevant to dyslexia. Given previous findings suggesting that these individuals are seeking similar types of support at university, I predicted that cases of dyspraxia in my sample would demonstrate a similar profile of impairment to cases of dyslexia. Conversely, I expected that adults with other developmental disorders such as autism and SpLD would be more similar to controls on these measures.

4.1.4 *Summary of chapter aims*

To summarise, the primary aims of this chapter were:

1. **Aim 1:** To provide a characterisation of the profile of impairments demonstrated by a sample of adults with developmental disorders across a range of standardised tests, which will be the subject of study of a further chapter investigating the relationship between language abilities and lateralisation (*Chapter 5*).
2. **Aim 2:** To consider which measures discriminated between the current group of high functioning dyslexic participants and the rest of the sample, including controls and other developmental disorders.
3. **Aim 3:** To explore whether individuals with a diagnosis of dyspraxia tend to cluster together with individuals with a diagnosis of dyslexia within a data-driven clustering analysis using measures relevant to dyslexia.

4.2 Methods

4.2.1 *Participants*

Participants were 104 adult individuals (36 male), 67 of whom had a prior diagnosis of a developmental disorder (DD group, mean age = 24.16) and 37 of whom had no such diagnosis (control group, mean age = 24.54). Mean age did

not differ significantly between these groups ($t(86.87) = .339$, $p = .735$). Exclusion criteria for both groups were significant hearing loss, history of neurological disease, head injury or epilepsy. Inclusion in the DD group did not rely on the application of psychometric criteria from collected data, but simply on self-report from the participant of a formal diagnosis of a recognised developmental disorder (including dyslexia, dyspraxia and autism) or specific learning difficulties (SpLD). Most participants were current students enrolled on a university or college of higher education course (82 participants); only 8 participants (6 controls, 2 DD) had never been through higher education. This was therefore a high functioning sample. Recruitment was predominantly via disability services at local institutes of higher education for the DD group (including the University of Oxford, Oxford Brookes University, New Bucks University and City, University of London), and via an online participant recruitment system for the control group.

The DD group consisted of individuals with a range of diagnoses, the most dominant being dyslexia. Frequencies in different diagnosis groups are given in Figure 4.1. Individuals in the category 'SpLD' (specific learning difficulties) did not reach criteria for any specific diagnosis such as dyslexia, but had been diagnosed by their university disability advisory service as having specific processing difficulties relative to other aspects of cognition, such as slower audio-visual processing.

According to self-report questionnaire data from the 67 individuals in the DD group, 35 (52.2%) had received special help or learning support whilst at school or university. The majority of these participants (33) reported receiving learning support whilst at school, including attending extra sessions on reading, spelling, handwriting and organisation. Most of these were also continuing to receive

special help at university, most commonly in the form of extra time and the use of a laptop in exams. Only 7 (10.4%) of the DD group reported having received speech and language therapy. A family history of similar problems was reported in just over half this group, with 36 (53.7%) reporting they had a family member with a history of speech, language, reading or communication disorder. The majority of the participants received their diagnosis whilst at school (44 participants, 65.7%), with the rest being diagnosed at university (except for one participant diagnosed through work).

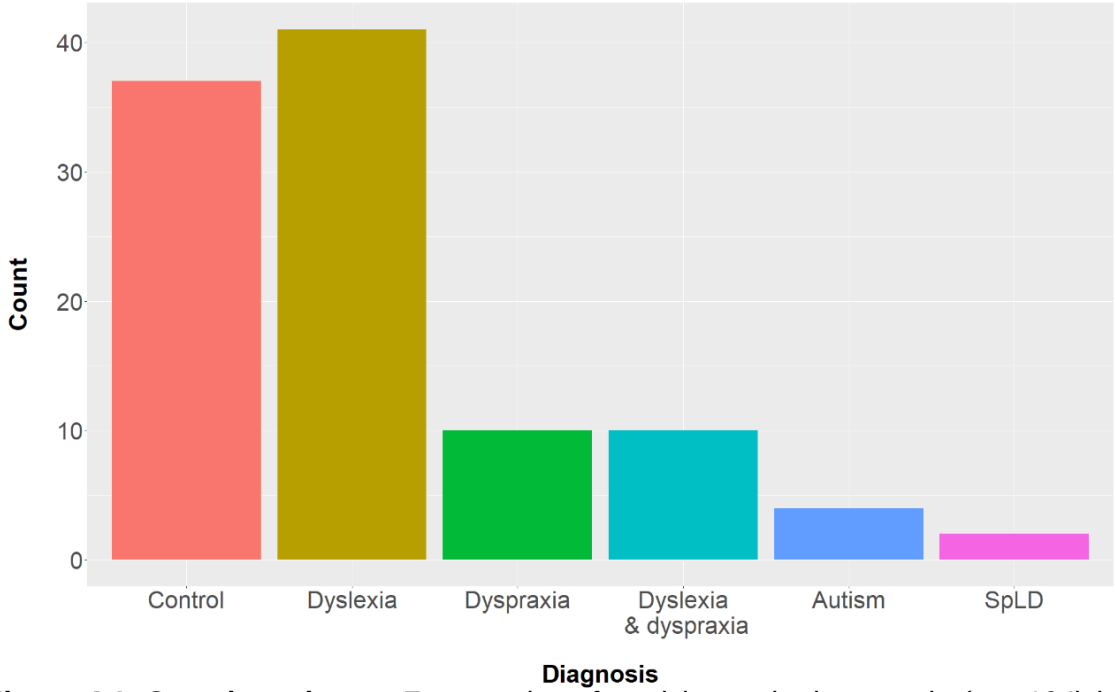


Figure 4.1: Sample make-up. *Frequencies of participants in the sample (n = 104) in each diagnosis group. SpLD = Specific Learning Difficulties.*

4.2.2 ***Procedure***

Data was collected as part of a larger study for which all participants completed an initial session involving administration of a language assessment battery. This session lasted up to two hours. Firstly, participants filled out a questionnaire asking for demographic information, and information on support received at school, history of speech and language therapy, and family history of speech,

reading or communication disorders. Three measures of handedness were administered; the Quantification of Hand Preference (QHP) task (Bishop, Ross, Daniels, & Bright, 1996); the Annett peg moving task (Annett, 1985); and the Edinburgh handedness inventory (Oldfield, 1971). A language assessment battery was then administered, consisting of a range of standardised tests designed to tap different aspects of language functioning. These are detailed below. A measure of non-verbal IQ was also obtained using tests one and two of the Cattell Culture fair test (Cattell & Cattell, 1973).

Data collected using these measures was stored on REDCap electronic data capture tools (Harris et al., 2009) hosted at Oxford University. REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies, providing: (1) an intuitive interface for validated data entry; (2) audit trails for tracking data manipulation and export procedures; (3) automated export procedures for seamless data downloads to common statistical packages; and (4) procedures for importing data from external sources.

4.2.3 *Language assessment battery*

The tests and their associated measures are listed in Table 4.1, along with the domain that each aims to test. A sub-set of 12 of these language measures were selected for the analysis presented in a later chapter (*Chapter 5*); these are highlighted in bold in Table 4.1. Each test is described in full in the following sections.

Table 4.1: Summary of measures from tests administered within the language assessment battery. *The 12 measures used in Chapter 5 are highlighted in bold (on next page).*

Instrument	Measure(s)	Domain being tested
Expression, Reception and Recall of Narrative Instrument (ERNNI) (Bishop, 2004)	1. Initial storytelling 2. Story recall **3. Mean length of utterance (MLU)** **4. Story comprehension**	Expressive language and story comprehension
Wechsler Adult Intelligence Scale (WAIS-IV) (Wechsler, 2008)	5. Digit span forwards 6. Digit span backwards **7. Digit span overall score**	Phonological short term memory Working memory
Wechsler Abbreviated Scale of Intelligence (WASI) (Wechsler, 1999)	**8. Vocabulary**	Verbal IQ
York Adult Assessment Battery- Revised (YAA-R) (Warmington, Stothard, & Snowling, 2013)	9. Rapid naming digits **10. Rapid naming objects**	Speed of processing
Test of Word Reading Efficiency (TOWRE) (Torgesen, Wagner, & Rashotte, 1999)	11. Word reading standard score 12. Non-word reading standard score **13. Overall reading standard score**	Single word recognition Phonological skills (decoding)
NEPSY- A developmental neuropsychological assessment (Korkman, Kirk, & Kemp, 1998)	**14. Oromotor sequences** **15. Non-word repetition**	Articulation/oromotor skills Phonological skills/auditory short term memory
Test of Adolescent and Adult Language (TOAL-4) (Hammill, Brown, Larson, & Wiederholt, 2007)	**16. Sentence combining**	Syntax/sentence production
Communication Checklist self-report (Bishop, Whitehouse, & Sharp, 2009)	**17. Language structure** **18. Pragmatics** **19. Social engagement**	Communication abilities (oral language skills, pragmatic skills and social engagement)
Cattell Culture Fair Test (Cattell & Cattell, 1973).	20. Tests 1 and 2 overall score	Non-verbal intelligence

4.2.3.1 Expression, Reception and Recall of Narrative Instrument (ERRNI)

The ERRNI (Bishop, 2004) required the participant to look through a set of pictures that tell a story and then narrate that story using the pictures. After an interval of between 15 to 30 minutes, the participant was asked to retell the story without the pictures. Initial storytelling and story recall were scored for content, and a combined measure of mean length of utterance calculated across the narratives. The participant was also asked some comprehension questions about the story, which tested both factual memory (e.g. What type of pet did the boy have?) and inferences about the mental state of characters (e.g. How did the boy feel when he found the doll?), to give a comprehension score for each participant. Raw scores were used for the main analyses of this chapter, however this test also provides standard scores based on a normative sample of 39 UK adults aged 17 to 64 years (mean age = 35.4). These standard scores were used for comparisons between this normative sample and the current sample (see *4.3.1 Aim 1: Statistical analyses*). This test reports a good level of internal reliability when used with a sample of typically developing and language impaired children aged 6 to 13 years (n = 75); specifically, moderate correlations were found between the two versions of the ERRNI (beach story versus fish story) which were all highly significant for the measures employed in the current study.

4.2.3.2 Wechsler Adult Intelligence Scale (WAIS) IV: Digit span

The WAIS-IV digit span test (Wechsler, 2008) required the participant to listen to and report back to the experimenter lists of numbers of increasing length. For the first set of numbers, participants had to recall the digits in the same order as given (forwards digit span) to test phonological working memory. A second set of numbers was then given which the participant was instructed to report in reverse

order (backwards digit span) to test auditory working memory. This gave separate scores for forwards and backwards digit span as well as a combined overall score. Although raw scores were used for most analyses, again standard scores based on the normative sample provided by the test were used for comparisons with the current sample. Norms were derived from a representative sample of 2200 US adults aged 16 to 90 years. The working memory composite made up of several sub-tests within the WAIS-IV including the digit span task is reported by the manual to have high reliability of .94.

4.2.3.3 Wechsler Abbreviated Scale of Intelligence (WASI): Vocabulary

The Vocabulary sub-test of the WASI (Wechsler, 1999) was administered to obtain a measure of vocabulary knowledge. This required participants to verbally describe and define words of increasing difficulty/obscurity, which were then scored for accuracy. Standardised age-normed scores were calculated using the test manual. These standard scores are based on a normative sample of 2245 US children and adults aged 6 to 89 years. For each age group, raw scores were converted into percentiles and then into standardised *T*-scores. The reliability of this sub-test was reported to be high within the adult age-band (aged 17 to 89 years) of this normative sample, with split-half reliability of .94 and test-retest reliability of .90.

4.2.3.4 York Adult Assessment Battery- Revised (YAA-R): Rapid naming

The rapid naming sub-tests of the YAA-R (Warmington et al., 2013) were administered to assess naming speed/speed of processing. Participants were required to name as fast as possible a series of written digits, and subsequently a series of pictures of objects. Naming rate (items per second) was calculated for digits and objects separately. Raw scores were used for all analyses, but data on

a normative sample provided by the test were used for comparison with the current sample. This normative sample comprised 106 UK university students aged 18 to 36 years, with no reported history of reading problems. The YAA-R does not provide information on reliability for this sub-test.

4.2.3.5 Test of Word Reading Efficiency (TOWRE)

The TOWRE (Torgesen et al., 1999) was administered to obtain measures of word and non-word reading ability. Participants were required to read a list of words/non-words as fast as they could within a time limit of 45 seconds. Standardised age-normed scores were then calculated for word and non-word reading separately, and combined to form an overall standard score. These standard scores provided by the test are based on a normative sample of 1507 US participants aged 6 to 24 years. For the current analyses, a small adjustment to the upper end of the distribution of these standardised scores was made so as to reduce ceiling effects. High internal reliability was reported for the 18-24 age-band within this normative sample, with correlations between scores from alternative forms ranging from .89 to .94 across different measures. The test also reports high test-rest reliability within a sample aged 19 years, with correlations between scores taken at two time points ranging from .82 to .96 across the different measures.

4.2.3.6 NEPSY- A developmental neuropsychological assessment

Two sub-tests of the NEPSY (Korkman et al., 1998) were administered; oromotor sequences and non-word repetition. The oromotor sequences test required participants to repeat each of a set of phrases out loud five times in a row; the number of failed attempts was recorded (e.g. omissions, distortions, substitutions of words) and an overall score calculated for the number of successful utterances.

This provides a measure of oromotor language skills. The non-word repetition task required participants to repeat out loud auditorily presented non-words. Accuracy of pronunciation of individual syllables was scored to give an overall accuracy score across words. This provides a measure of phonological skills and auditory short-term memory. Standard scores for these tasks are only provided by the manual based on a child sample, thus raw scores were used instead of standard scores. However, a paper by Barry, Yasin and Bishop (2007) collected normative data using these tasks with a sample of 33 UK adults; the mean and standard deviation of the scores of this group will be used to derive norms for comparison with the current sample (see *4.3.1 Aim 1: Statistical analyses)*. Based on the child sample reported on by the test (aged 3 to 12 years), moderate levels of test-rest reliability were reported of .67 for the non-word repetition sub-test, and 73% consistency of classification for the oromotor sequences sub-test.

4.2.3.7 Test of Adolescent and Adult Language: Sentence assembly

The sentence assembly sub-test of the TOAL-4 (Hammill et al., 2007) was administered to obtain a measure of syntactic abilities. The participant was given sets of short sentences written on cards, and required to orally combine each set into a single sentence that preserved the meaning of the set of sentences. This task was self-paced, and participants were able to view the written sentences throughout. Produced sentences were marked for their grammatically and content, with full marks awarded for a well formed sentence that included all ideas expressed in the set of sentences. Standard scores are provided by the test manual based on a normative sample of 1671 US participants aged 12 to 24 years. Raw scores were used for the main analyses, but standard scores used to compare the developmental disorder group to this normative group. The TOAL

reports a high level of reliability for this sub-test, with internal consistency reliability of .89 (Cronbach's alpha) averaged across age bands, and a test-rest reliability coefficient of .93 for the 15 to 18 year age band.

4.2.3.8 Communication Checklist self-report (CC-SR)

The CC-SR (Bishop et al., 2009) was administered to obtain self-report measures of each participant's language and communication abilities. This is a questionnaire made up of a set of statements for which the participant must report how often each happens to them using a coded scale. These statements form three sub-scales of language structure (oral language ability), pragmatic skills and social engagement. Age-adjusted standard scores were calculated using the automated scorer provided with the test. These standard scores are based on a normative sample of 481 UK participants between the ages of 10 and 89. Z scores are first computed for three broad age bands (10 to 11 years, 12 to 25 years and 26 to 89 years), which are then converted into scaled scores. The average internal reliability coefficient across the three sub-scales was reported as .93 with this normative sample.

4.2.4 ***Normality of the measures***

Distributions of scores on the 12 language measures used in the analysis of Chapter 5 are given in Figure 4.2 for the developmental and control groups separately. Examination of the distributions of scores across the language measures in the groups revealed multiple cases of deviations from normality, confirmed in Shapiro-Wilks tests. As well as skew, several of the language measures demonstrate signs of bimodality (see plots for ERRNI comprehension, TOWRE overall standard score, and CCSR measures). While transformations such as the Box-Cox transform are able to deal with skewness, they are not

successful at inducing normality in bimodally distributed measures. Therefore, when comparing groups on language measures for which at least one of the groups demonstrated a non-normal distribution of scores, non-parametric tests were used.

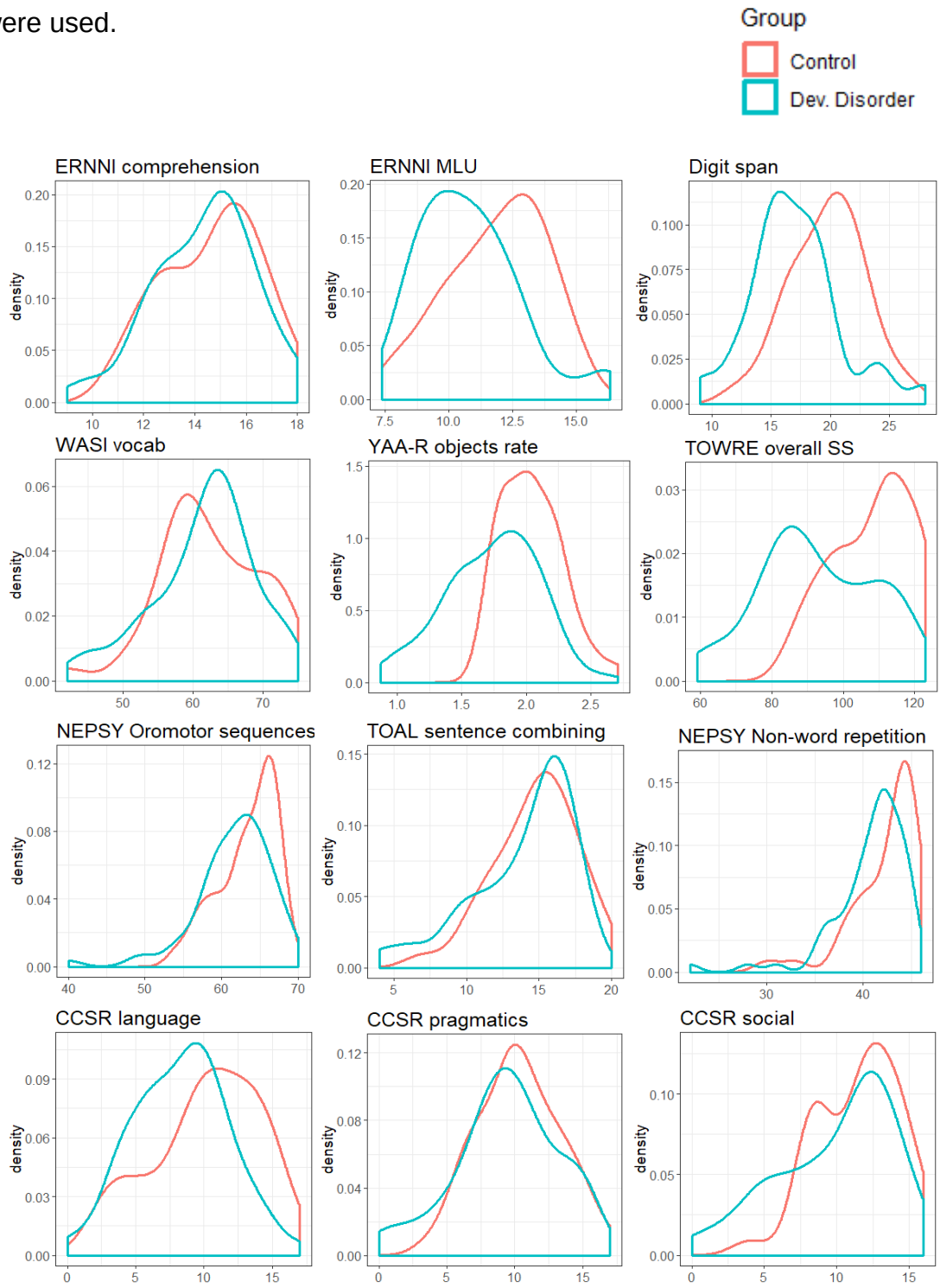


Figure 4.2: Language measure distributions. *Distributions of scores in the control and developmental disorder groups on the 12 language measures.*

All analyses were conducted using R software (R Core Team, 2019). An R markdown script that runs all analyses reported here can be found on the OSF page for this project (https://osf.io/bwsha/). In order to aid clarity and flow, the following sections will describe statistical analyses and report results for each of the three aims in turn.

4.3 Aim 1: Characterisation of the profile of impairments in a sample of adults with developmental disorders

4.3.1 *Statistical analyses*

In preparation for analyses presented in a later chapter of this thesis investigating language laterality patterns in this sample (*Chapter 5*), I first aimed to investigate the extent and nature of language impairments present in the DD group. As an initial step, a non-parametric multivariate test was used to test for a significant difference between the performance of the DD and control groups on the 12 selected language measures as a set. Use of a non-parametric test was necessary due to the non-normality of the language measures (see *4.2.4 Normality of the measures*). Univariate non-parametric tests use the metrics of signs and ranks which order the data based on a single measure. In the multivariate case where one wishes to compare across multiple measures simultaneously, there is no natural ordering of the observations since each observation is now a vector with length p (where p corresponds to the number of measures); thus ordering based on one measure will conflict with ordering based on another measure. Instead, multivariate spatial signs and spatial ranks are used to provide an ordering of the observations which takes into account all measures. In *p*-dimensional space, the multivariate spatial signed rank provides a unique value for each observation which is based on its Euclidian distance.

These transformed values can be used to calculate the corresponding test statistic, Q^2, which follows a chi-squared distribution and thus can be compared with the chi-squared critical value to test for significance. This test can therefore be considered as a multivariate extension of the Wilcoxon-Mann-Whitney test, but with altered calculation of signs and ranks to reflect the multidimensionality of the data. This non-parametric multivariate test was run using the R package MNM (Nordhausen & Oja, 2015) to compare performance of the two groups (DD versus control) on the 12 language measures.

I predicted that this comparison would be significant, indicating that the DD group demonstrated significantly worse performance than the control group on the language measures as a set. This analysis was important for the later chapter (*Chapter 5*) in order to demonstrate that these language measures were sensitive enough to distinguish between the two diagnosis groups.

In addition to comparing the DD group to the current control sample, standardised scores were also used to compare the performance of the DD and control groups to that of the normative samples provided by the tests. This was done to demonstrate the extent to which the scores of the DD group were in the impaired range relative to general population norms. To do this, the percentage of participants scoring below a specified threshold on each measure was calculated for DD and control groups separately. A summary of the scores, normative samples and cut-offs used for each measure is given in Table 4.2. Where standard scores were provided by the test on an appropriate normative sample, the percentage of participants with a standard score that was one standard deviation below that of the normative sample mean was calculated for each measure. Otherwise, alternative cut-offs were defined as specified in the table.

Table 4.2: Summary of procedure for normative comparisons. *Details of standardisation procedures for comparison of DD and control groups with normative samples across the 12 language measures (continued on next page).*

Measure	Scores used	Normative sample description	Cut-offs used
ERRNI comprehension	Standard scores	39 UK adults aged 17 to 64 years. Mean score of 102.12 (*SD* = 16.09)	1 SD below normative sample mean
ERRNI mean length of utterance	Standard scores	39 UK adults aged 17 to 64 years, mean score of 98.02 (*SD* = 16.93)	1 SD below normative sample mean
WAIS-IV Digit span	Standard scores	2200 US adults aged 16 to 90 years.	Score of 7 or below (described as 'below average' by Sattler & Dumont, 2004)
WASI vocabulary	T-scores	2245 US children and adults aged 6 to 89 years. Mean score of 50 (*SD* = 10).	1 SD below normative sample mean
YAA-R object rapid naming	Raw scores	106 UK university students aged 18 to 36 years. Mean score of 1.88 (*SD* = .35)	1 SD below normative sample mean
TOWRE overall	Standard scores	1507 US adults aged 12 to 24 years. Mean score of 100 (*SD* = 15)	1 SD below normative sample mean
NEPSY oromotor sequences	Raw scores	33 UK adults aged 34 to 56 years (from Barry et al., 2007). Mean score of 64 (*SD* = 4.6)	1 SD below normative sample mean

NEPSY non-word repetition	Raw scores	33 UK adults aged 34 to 56 years (from Barry et al., 2007). Mean score of 41.0 (*SD* = 3.9)	1 SD below normative sample mean
TOAL sentence assembly	Standard scores	1671 US participants aged 12-24 years. Mean score of 10 (*SD* = 3)	1 SD below normative sample mean
CCSR language structure	Scaled scores	481 UK participants aged 10 to 89 years. Mean score of 10 (*SD* = 3)	1 SD below normative sample mean
CCSR pragmatics	Scaled scores	481 UK participants aged 10 to 89 years. Mean score of 10 (*SD* = 3)	1 SD below normative sample mean
CCSR social engagement	Scaled scores	481 UK participants aged 10 to 89 years. Mean score of 10 (*SD* = 3)	1 SD below normative sample mean

4.3.2 ***Results***

Standardised scores for the DD and control groups across the 12 measures are plotted in Figure 4.3. The non-parametric multivariate test found a significant difference between developmental disorder and control groups across the 12 language measures (Q^2 = 34.06, *df* = 12, *p* < .001). This indicates that the language measures were sensitive enough to distinguish between the developmental disorder and control groups.

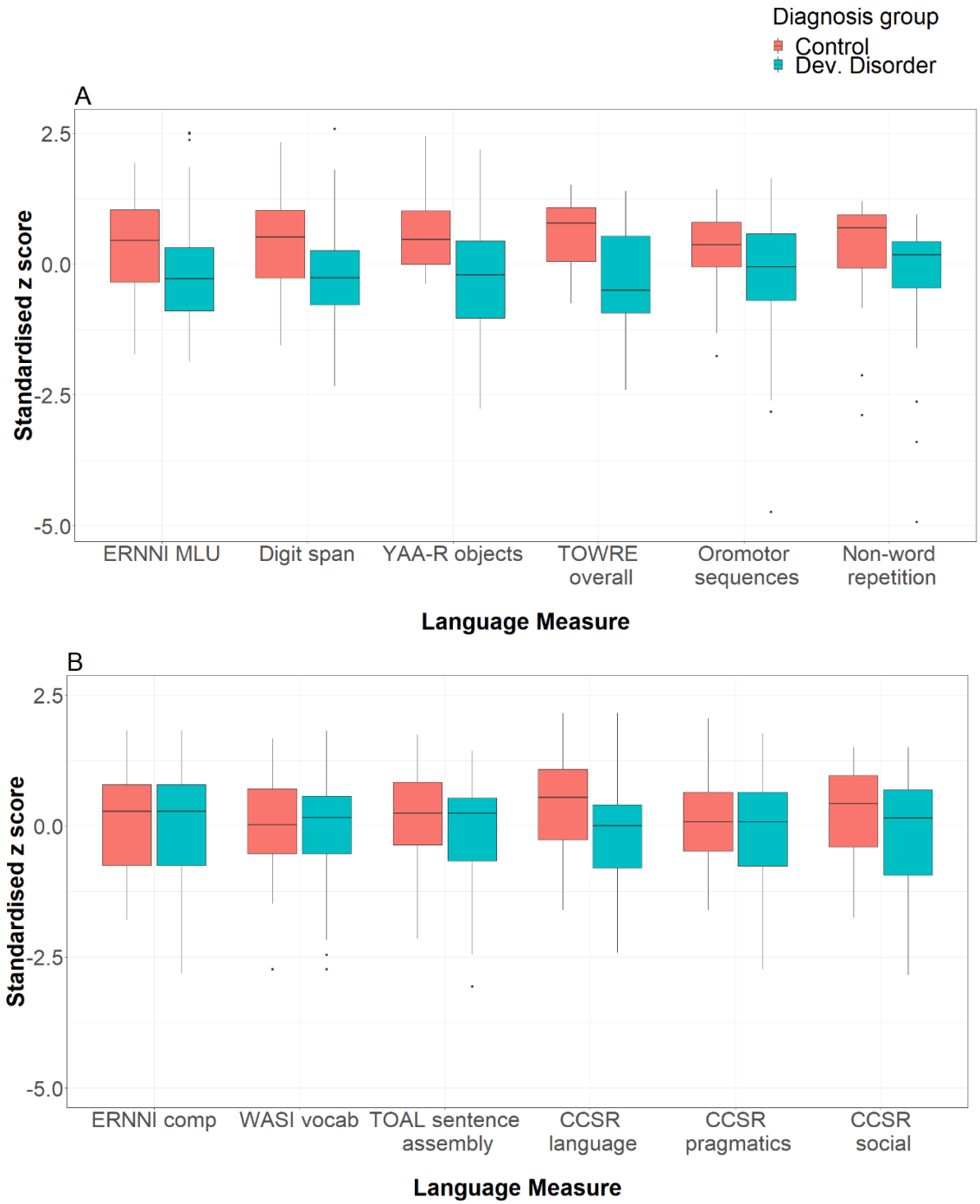


Figure 4.3: Standardised language scores. *Boxplots of standardised scores on each of the 12 language measures in the developmental disorder and control groups. The first six measures are shown in panel A, and the last six measures in panel B. Dark lines indicate medians for each group, dots indicate outliers.*

The percentages of participants in the DD and control groups scoring below cut-offs derived from normative samples (see Table 4.2) are given in Table 4.3.

These demonstrate that across measures, a low to moderate proportion of participants within the DD group are impaired relative to the general population; in particular, relatively high percentages are seen for YAA-R rapid naming, TOWRE word and non-word reading, NEPSY oromotor sequences, TOAL sentence assembly and CCSR language structure measures. For the control group percentages are generally very low, with notable exceptions being TOAL sentence assembly, CCSR language structure and NEPSY oromotor sequences. This is likely attributable to demographic differences between this sample and the normative samples employed by the tests, for example in age and nationality. Comparing between groups, it is particularly apparent that the high percentage of DD participants scoring below normative cut-offs on the TOAL sentence assembly test should not be interpreted as indicating a particular weakness of this group for this aspect of language processing.

Table 4.3: Comparison of DD and control groups with normative samples. *Percentages of DD and control groups scoring below normative cut-offs across the 12 language measures. DD = developmental disorder.*

Measure	Percentage DD group below cut-off	Percentage control group below cut-off
ERRNI comprehension	5.97	5.41
ERRNI MLU	2.99	2.70
Digit span	8.96	2.70
WASI vocabulary	0.00	0.00
YAA-R objects rapid naming rate	32.84	0.00
TOWRE overall	34.33	0.00
NEPSY oromotor sequences	26.87	18.92
NEPSY non-word repetition	14.93	5.40
TOAL sentence assembly	35.82	29.73
CCSR language structure	32.84	21.62
CCSR pragmatics	16.42	10.81
CCSR social engagement	22.39	2.70

4.4 Aim 2: Investigation of which measures were most sensitive for discriminating dyslexic from non-dyslexic participants

4.4.1 *Statistical analyses*

To investigate which measures were sensitive to a diagnosis of dyslexia as opposed to other diagnoses or no diagnosis for the current high functioning sample, the sample was split into those with a diagnosis of dyslexia (either on its own or comorbid with other disorders, n = 51), and those without a diagnosis of dyslexia (including both controls and other disorders, n = 53). To begin with, I took a traditional approach by obtaining effect sizes (Cohen's d) for differences between the two groups for each measure. These were obtained using either parametric t-tests or non-parametric Mann-Whitney-U tests, depending on the normality of each measure within each group (see Figure 4.2). In the non-parametric case, Cohen's d was calculated from the *r* effect size estimate using an online converter (https://www.psychometrica.de/effect_size.html). Despite this being an established approach in the previous literature, the application of multiple univariate tests to this data is sub-optimal, both because of the inflation of type one error but also because the Bonferroni correction that aims to correct for this is likely too stringent. Therefore, this analysis was followed up using the same non-parametric multivariate test (Nordhausen & Oja, 2015) as described above for the comparison of DD and control groups. For this analysis, bootstrapping (Efron, 1979) was used to obtain 95% confidence intervals for mean differences between groups on each of the measures, in order to ascertain on which measures the two groups significantly differed, whilst controlling the type 1 error rate.

In line with previous research, it was predicted that the two groups would differ on measures of word reading, phonological skills (non-word reading, non-word repetition), phonological short term memory/working memory (digit span), speed of processing (rapid naming), oral language (oromotor sequences, CCSR language structure, ERRNI MLU) and sentence assembly (TOAL). In contrast, it was predicted that the two groups would not differ on measures of vocabulary (WASI), comprehension (ERRNI comprehension) and non-verbal IQ (Cattell).

Following on from this investigation of group differences, a logistic regression analysis was conducted to investigate which measures were significant predictors of a dyslexia diagnosis for our high functioning participants. Measures were chosen based on which demonstrated significant differences between groups in the above outlined analyses. The multicollinearity of these variables was checked using Spearman's correlations. A two stage approach was then taken, in which a series of logistic regression models were successively added to one of two null models. The first null model included TOWRE word reading, while the null model used in the second phase included TOWRE non-word reading. The measures added to each null model in a series of comparison models for each phase are given in Table 4.4. Likelihood ratio tests were used to test the benefit of the addition of each variable compared to the null model in terms of an increase in the goodness of fit of the model for discriminating dyslexic from non-dyslexic cases. For the first phase, it was predicted that measures of phonological skills, working memory and language skills would all improve prediction relative to a measure of single word reading alone. For the second phase, it was predicted that the addition of variables relating to working memory and oral language skills would further improve prediction over phonological skills alone.

Table 4.4: Logistic regression models. *Summary of models used in logistic regression analyses, split into phase 1 and phase 2 which each began with a different null model.*

Model	Measures included
Phase 1	
Null model	TOWRE words
Model 2	TOWRE words TOWRE non-words
Model 3	TOWRE words Digit span
Model 4	TOWRE words NEPSY oromotor sequences
Phase 2	
Null model	TOWRE non-words
Model 2	TOWRE non-words Digit span
Model 3	TOWRE non-words NEPSY oromotor sequences

4.4.2 ***Results***

Effect sizes and p values for differences between dyslexic and non-dyslexic groups are given for the full set of language measures in Table 4.5. Standardised scores across these 14 measures are given in Figure 4.4 for the two groups. Multiple comparisons suggested significant group differences across measures of literacy, phonological processing, processing skills and oral language skills. These differences were more optimally investigated using a non-parametric multivariate analysis approach. As expected, this found that the two groups differed across the measures as a set, $Q^2 = 41.18$, $df = 12$, $p < .001$. Bootstrapped 95% confidence intervals revealed significant differences between groups on TOWRE word and non-word reading, YAA-R object rapid naming, digit span, NEPSY oromotor sequences and non-word repetition. Conversely, no significant differences were found for Cattell culture fair non-verbal IQ, WASI vocabulary or ERRNI comprehension.

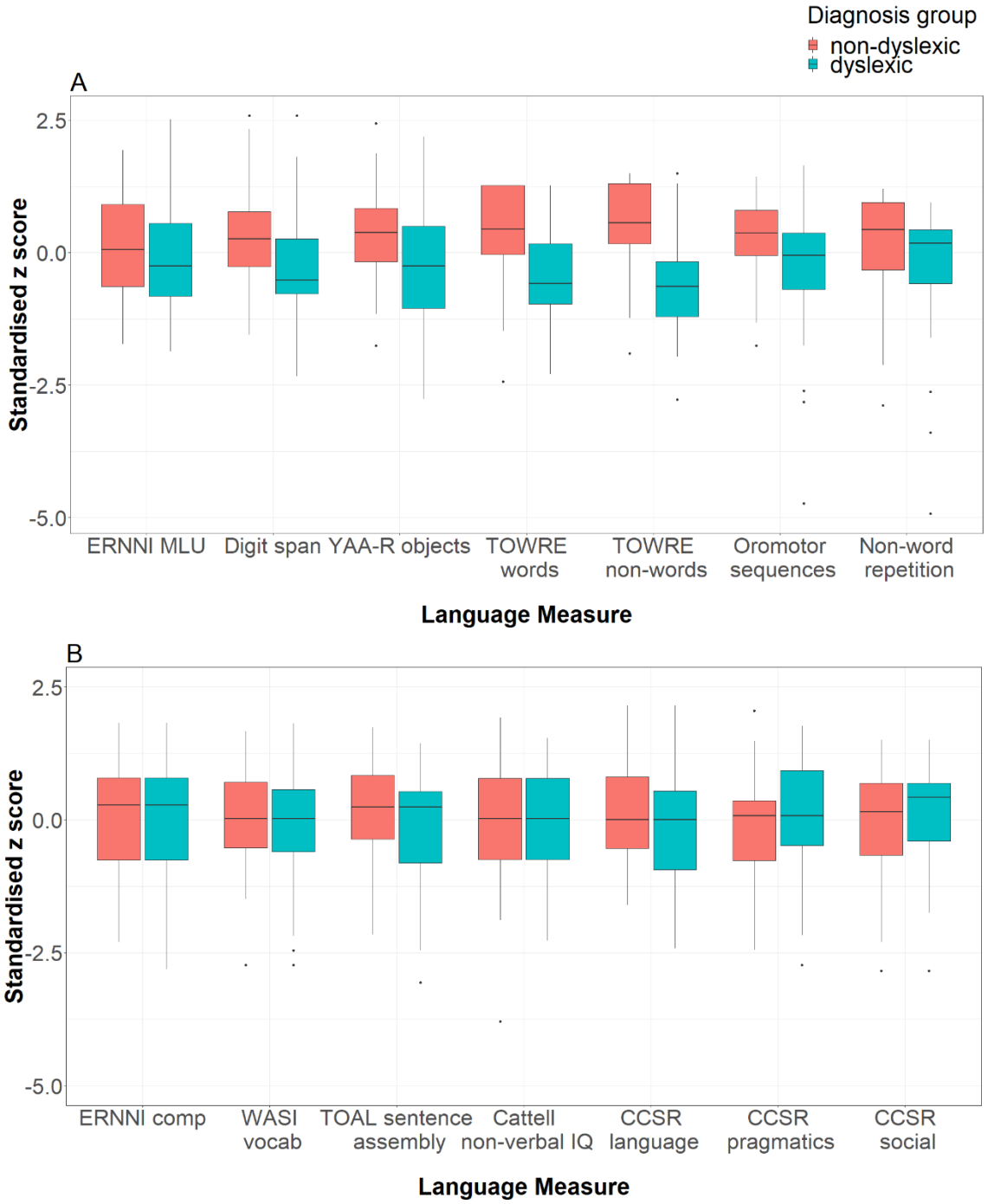


Figure 4.4: Standardised language scores in dyslexic and non-dyslexic groups. *Boxplots of standardised scores on each of the 14 language measures in the dyslexic and non-dyslexic groups. The first seven measures are shown in panel A, and the last seven measures in panel B. Dark lines indicate medians for each group, dots indicate outliers.*

Table 4.5: Performance of dyslexic and non-dyslexic groups. *Effect sizes (Cohen's d) and p values for differences between participants with and without a diagnosis of dyslexia. The sign of effect sizes has been inverted such that a decrease in performance in dyslexics versus controls corresponds to a positive effect size. Measures that were non-normal and therefore compared by means of a non-parametric test are marked with [1]. Significant p values are highlighted in bold.*

Measure	Mean (SD) dyslexics	Mean (SD) non-dyslexics	Effect size (*d*)	*p*
Literacy				
TOWRE words[1]	91.00 (13.84)	101.87 (13.346)	0.849	**<.001**
Phonological processing				
TOWRE non-words[1]	92.04 (13.488)	108.66 (11.484)	1.402	**<.001**
NEPSY non-word repetition[1]	40.43 (4.41)	42.11 (3.21)	0.474	**.018**
Processing skills				
Digit span total score	16.76 (3.85)	19.19 (3.52)	0.657	**.001**
YAA-R objects rapid naming rate	1.73 (0.38)	1.96 (0.28)	0.705	**<.001**
Oral language skills				
ERRNI mean length of utterance[1]	10.98 (2.14)	11.37 (1.95)	0.256	.197
NEPSY oromotor sequences[1]	61.04 (5.39)	63.42 (3.62)	0.505	**.012**
CCSR language	8.43 (3.85)	9.51 (3.54)	0.292	.140
Vocabulary and comprehension				
WASI vocabulary	61.14 (7.92)	62.51 (6.59)	0.189	.340
ERRNI comprehension	14.27 (2.00)	14.64 (1.88)	0.189	.338
Syntactic processing				
TOAL sentence assembly[1]	13.65 (3.74)	14.74 (2.83)	0.236	.230
Social and pragmatic skills				
CCSR pragmatics[1]	9.92 (3.93)	9.51 (3.18)	0.178	.365
CCSR social[1]	10.75 (3.48)	10.17 (3.87)	0.139	.479
General intelligence				
Cattell non-verbal IQ[1]	19.96 (2.37)	19.94 (2.87)	0.011	.958

A correlation heat map for relationships between these six significant measures is given in Figure 4.5. These generally demonstrate modest correlations, suggesting multicollinearity was not an issue for the subsequent logistic regression analysis.

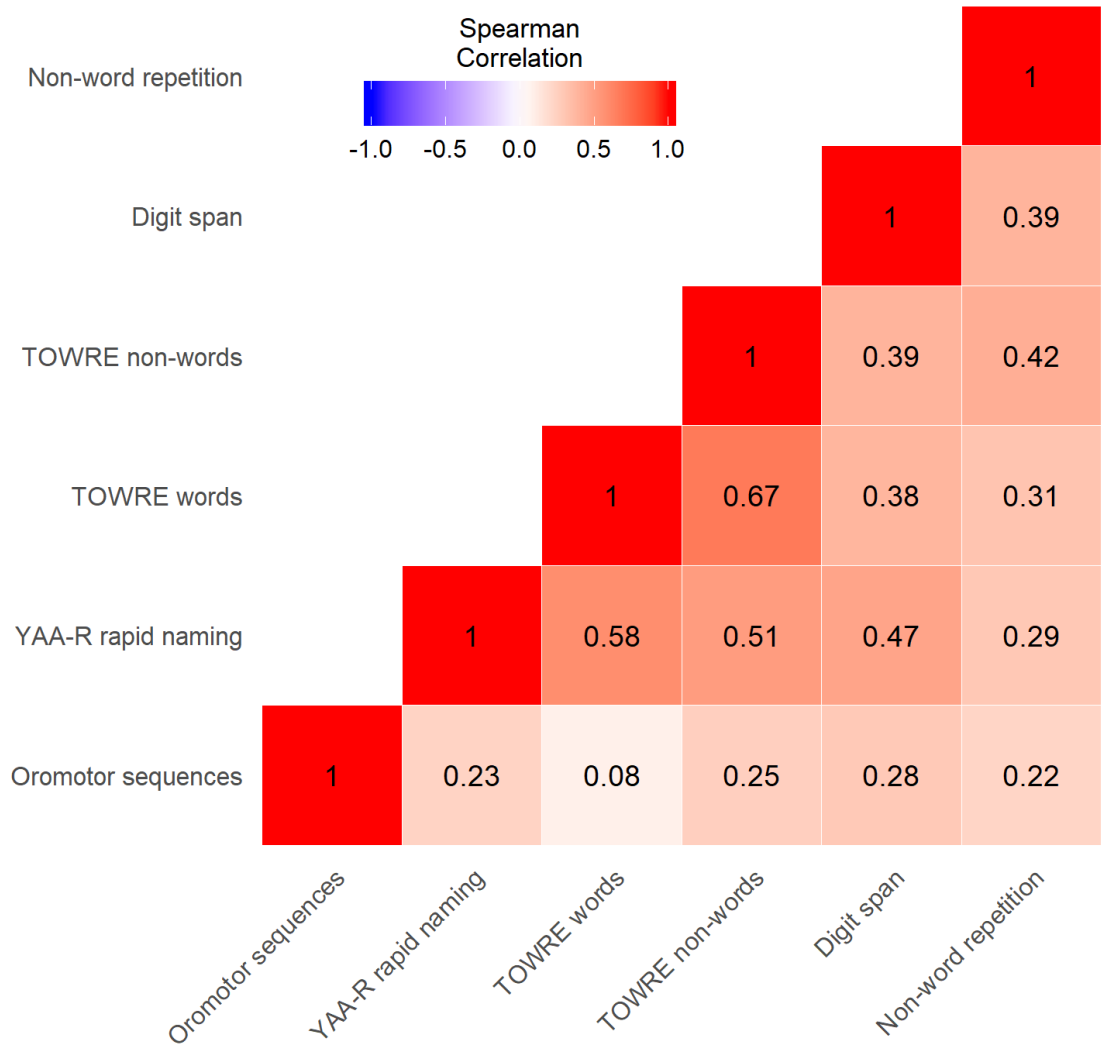


Figure 4.5: Correlations between measures. *Correlation heat map for relationships between the six measures that differed significantly between dyslexic and non-dyslexic groups on a multivariate test. Spearman's correlations are shown.*

To determine which of these measures were optimally sensitive for discriminating between the current high functioning dyslexic and non-dyslexic individuals, these significant variables from the above analysis were used in a series of logistic

regressions. Firstly, a null model including only TOWRE word reading scores was used to test for the additional benefit of variables relevant to phonological skills, processing skills and oral language skills using a likelihood ratio approach. The null model found that TOWRE word reading was a significant predictor of dyslexia (β = -0.057, p < .001, AIC = 132.79). In comparison to this null model, the addition of TOWRE non-word reading (AIC = 113.15), digit span (AIC = 130.47) and oromotor sequences (AIC = 129.18) all significantly improved model fit when added one at a time (see Tables 4.6 and 4.7).

Table 4.6: Phase 1. *Summary of logistic regression models predicting dyslexia diagnosis from different variables starting with a null model of single word reading.*

Model	Measures	β Coefficient	AIC
Null model	TOWRE words	-0.0574**	132.79
Model 2	TOWRE words	0.002	113.15
	TOWRE non-words	-0.101**	
Model 3	TOWRE words	-0.048*	130.47
	Digit span	-0.128*	
Model 4	TOWRE words	-0.056**	129.18
	NEPSY oromotor sequences	-0.126*	

** p < .001, * p < .05

Table 4.7: Likelihood ratio tests phase 1. *Results of likelihood ratio tests comparing logistic regression models adding phonological skills, working memory and oral language skills to a null model of single word reading.*

Model	Residual Df	Residual Deviance	Df	Deviance	p
Null model: TOWRE words	102	128.79			
Model 2: TOWRE words TOWRE non-words	101	107.15	1	21.64	< .001
Model 3: TOWRE words Digit span	101	124.47	1	4.32	.038
Model 4: TOWRE words Oromotor sequences	101	123.18	1	5.62	.018

In a second stage, TOWRE non-word reading was used as a null model, in order to test for the additional benefit of incorporating digit span and oromotor sequences into the model using likelihood ratio tests. TOWRE non-word reading alone was a significant predictor (β = -0.098, p < .001, AIC = 111.16). However, the addition of these extra variables did not provide better model fit over and above that provided by TOWRE non-word reading alone (see Tables 4.8 and 4.9). This suggests that for the current high functioning adult sample, incorporating working memory and language skills as predictor variables improved dyslexia classification over and above that achieved using only a single word reading measure, but not that achieved using a single measure of phonological skills (non-word reading).

Table 4.8: Phase 2. *Summary of logistic regression models predicting dyslexia diagnosis from different variables starting with a null model of non-word reading.*

Model	Measures	β Coefficient	AIC
Null model	TOWRE non-words	-0.098**	111.16
Model 2	TOWRE non-words Digit span	-0.093** -0.095	111.37
Model 3	TOWRE non-words NEPSY oromotor sequences	-0.096** -0.079	111.38

** $p < .001$, * $p < .05$

Table 4.9: Likelihood ratio tests phase 2. *Results of likelihood ratio tests comparing logistic regression models adding working memory and oral language skills to a null model of phonological skills (non-word reading).*

Model	Residual Df	Residual Deviance	Df	Deviance	p
Null Model: TOWRE non-words	102	107.16			
Model 2: TOWRE non-words Digit span	101	105.37	1	1.79	.181
Model 3: TOWRE non-words Oromotor sequences	101	105.38	1	1.78	.182

To better understand why the addition of other variables improves model fit when starting with TOWRE word reading but not TOWRE non-word reading, the relationship between scores on these two measures is plotted in Figure 4.6. This

demonstrates marked ceiling effects in TOWRE word reading, which may explain why there is more variance to be explained by additional variables in the regression model. Conversely, TOWRE non-word reading is not characterised by strong ceiling effects, and so may be able to explain more variance.

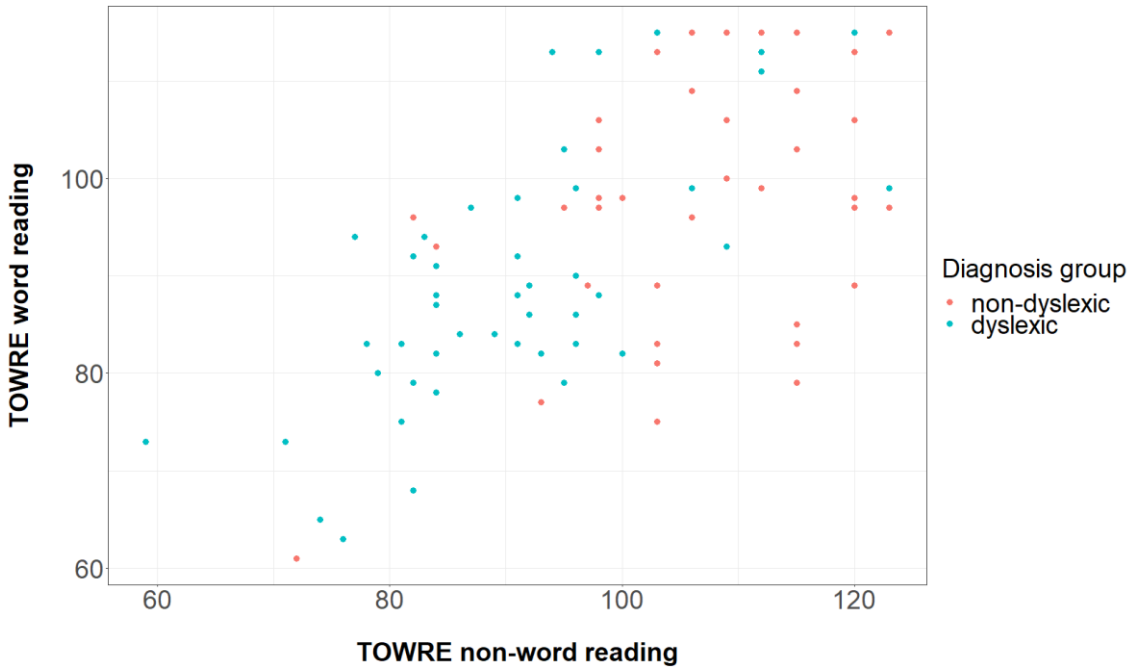


Figure 4.6: TOWRE scores. *Standard scores on the word reading and non-word reading sub-tests of the Test of Word Reading Efficiency (TOWRE) in dyslexic and non-dyslexic groups.*

4.5 Aim 3: Correspondence between diagnostic labels and data-driven sub-groups of participants

4.5.1 *Statistical analyses*

The aim of this analysis was to derive data-driven sub-groups from the language measures, and to examine whether these corresponded to diagnosis labels; in particular, I was interested in whether cases of dyspraxia within this high functioning sample tended to group together with the cases of dyslexia. To obtain data-driven sub-groups, I used an unsupervised machine learning (UML)

approach on language measures that emerged as significant when comparing dyslexic and non-dyslexic groups in *Aim 2*. The selected variables were TOWRE word reading, TOWRE non-word reading, digit span, YAA-R rapid naming objects and NEPSY oromotor sequences. NEPSY non-word repetition was not included on the basis that TOWRE non-word reading already provided a sensitive measure of phonological skills. For the current chosen method of clustering, interpretation of clusters can become difficult when including too many input variables, known as 'the curse of dimensionality' (Hastie, Tibshirani, & Friedman, 2001); five variables was considered an appropriate number for this sample size.

The UML approach seeks to find structure in unlabelled input data; that is, the groups are not pre-specified but derived by looking at similarities between cases in order to define homogenous sub-groups. A k-means clustering analysis was used to group individuals or 'observations'. K-means clustering aims to partition observations into a fixed number of k homogenous clusters. Each observation is considered as a point in multi-dimensional space. Initially, k number of points in this space are chosen to represent the centre of clusters, and observations assigned to their nearest cluster. Cluster centres are then re-calculated based on their cluster members, and observations reassigned to clusters based on these new centres. This process is repeatedly iteratively until cluster memberships stabilise. This was implemented using the R function 'kmeans' within the package 'stats', which selects the most optimal partitioning of cases based on which yields the smallest total within-cluster sum of squares. To determine the number of clusters (k) to specify, the k-means clustering process was iteratively repeated over different values of k and the average within-cluster sum of squares calculated. The optimal value of k was then chosen by identifying the 'elbow'

within a scree plot of within-cluster sum of squares at each value of k. It was predicted that the optimal solution would be a two-cluster grouping, with one cluster containing an overrepresentation of cases of dyslexia and dyspraxia, and the other containing mostly controls and the other disorders (autism and SpLD).

4.5.2 *Results*

As indicated by the scree plot shown in Figure 4.7, the optimal solution was a two-cluster grouping. K-means clustering analysis was then carried out with two clusters specified. The make-up of the two clusters derived by this approach in terms of diagnoses of the grouped cases is given in Figure 4.8. Cluster 1 contained 40 cases, most of which had a dyslexia diagnosis (27 dyslexic, 5 dyslexia and dyspraxia), with five with a dyspraxia diagnosis and three controls. Cluster 2 contained 64 cases, including most of the controls (34) and a selection from the other diagnosis categories (including 19 with a dyslexia diagnosis and 5 dyspraxics). As can be seen in Figure 4.8, individuals with a diagnosis of dyspraxia are fairly evenly split between the two clusters. This suggests that some adults with dyspraxia within the current high functioning sample were more similar to individuals with a dyslexia diagnosis than to controls in terms of their profile of abilities on these measures relevant to dyslexia, while others were more similar to controls. Conversely, all participants with diagnoses of autism and SpLD clustered with the controls, indicating that they did not have a dyslexic-like profile on these measures.

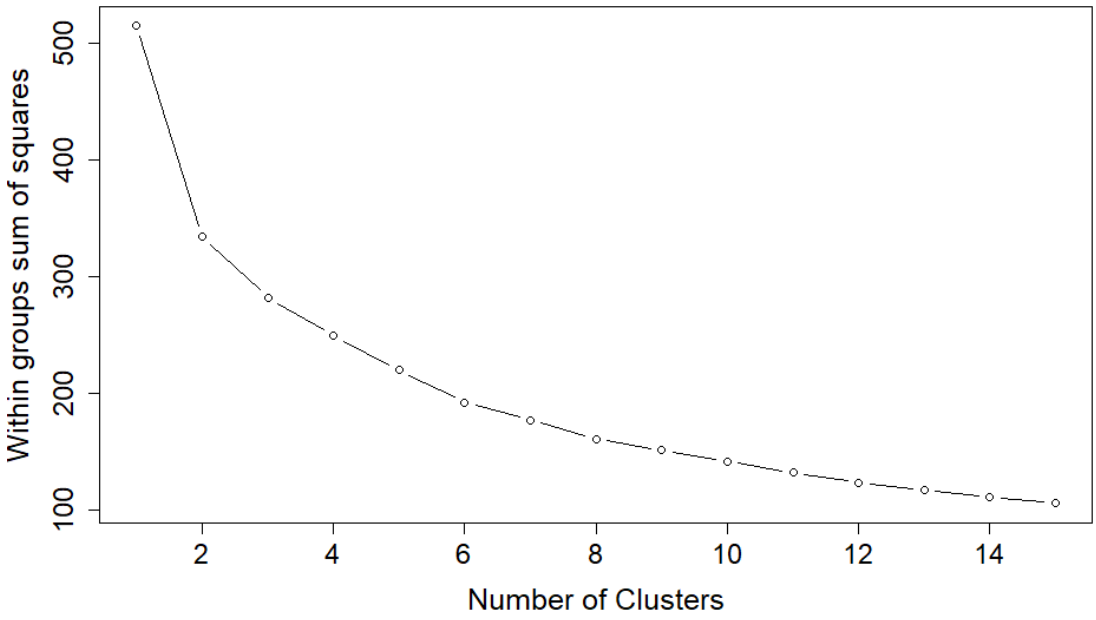


Figure 4.7: Scree plot for number of clusters. *Plot shows within-cluster sum of squares calculated across a range of values of k (number of clusters). An elbow is seen at a value of 2, indicating a 2 cluster solution is optimal.*

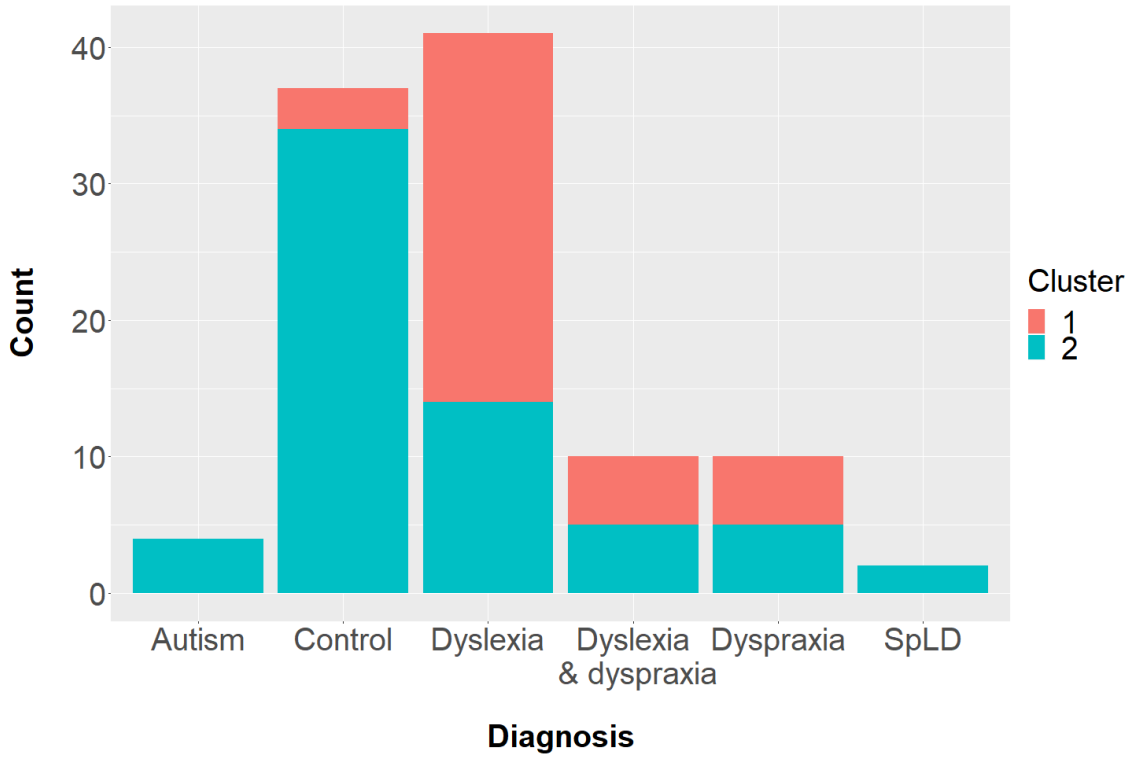


Figure 4.8: Cluster composition. *Make-up of the sample grouped into the two clusters by the k-means clustering analysis. Cluster 1 is predominantly made up of individuals with a dyslexia diagnosis, with a small number of dyspraxics and controls. Cluster 2*

contains the majority of the control participants, but also a number of individuals with a dyslexia diagnosis (19 in total) as well as 5 dyspraxics. SpLD = Specific Learning Difficulty.

4.6 General Discussion

This study aimed to examine the profile of impairment in a sample of high functioning adults with a history of developmental disorders. It was found that such adults demonstrated impaired functioning relative to controls across a range of tests of language, memory and processing skills. When considering specifically those participants with a diagnosis of dyslexia, impairments were found on measures of single word reading, phonological skills, processing speed, working memory and oral language skills. Conversely, spared functioning was seen for measures of non-verbal intelligence and vocabulary in line with previous research (e.g. Hatcher, Snowling, & Griffiths, 2002), as well as for measures of syntactic structure building and story comprehension (non-reading based). However, a logistic regression analysis found that phonological skills (non-word reading) alone were superior in discriminating the current highly functioning dyslexic individuals from non-dyslexic individuals, with additional variables of working memory and oral language failing to significantly improve prediction. Finally, a data-driven clustering analysis based on those measures most relevant to dyslexia yielded two clusters; one which predominantly included dyslexic individuals, and another which included the vast majority of control participants. However, dyslexic participants were still relatively split between the two clusters, as were cases of dyspraxia. Thus, in this analysis of an unusual high functioning sample based on a small selection of measures relevant to dyslexia, data-driven clusters did not neatly correspond with diagnostic labels.

4.6.1 *Aim 1: Characterisation of the profile of impairments in a sample of adults with developmental disorders*

In order to ascertain whether the current sample of adults with developmental disorders was impaired on the measures employed in the assessment battery, a multivariate test was used to compare DD participants with control participants across the measures as a set. This demonstrated that even in a high functioning adult sample (the majority of which were university students), a prior diagnosis of a developmental disorder was still associated with poorer language and memory skills relative to a comparable sample of university educated adults without such disorders. Comparisons with test standardization norms indicated that a moderate proportion of the DD group further demonstrated language skills in the impaired range relative to the general population.

For the purposes of this thesis, this demonstration was important for analyses presented in a later chapter (*Chapter 5*) examining the relationship between language abilities and language lateralisation within this sample. In this later chapter, it was predicted that individuals within this sample showing an inconsistent pattern of lateralisation across multiple language tasks would demonstrate impaired language skills relative to individuals with consistent lateralisation. Thus, it was important for this analysis to demonstrate that the sample contained a wide range of language abilities, and that the measures used were sensitive to the types of impairments present in the sample. A failure to demonstrate a significant difference between DD and control groups on these measures would have called into question the validity of the choice of sample or choice of measures for this later comparison of abilities between laterality groups.

However, it is notable that a high proportion of this group nevertheless demonstrated skills within the normal range when compared to standardised test norms based on more representative population samples. This confirms that the current sample was indeed a highly functioning sample compared to the general population, and was not severely impaired in terms of everyday functioning. Nevertheless, the comparison with the present more well-matched control group demonstrates that these adults with developmental disorders still show persisting areas of weakness relative to their non-disordered counterparts. This refutes the idea that success in entering higher education is associated with complete resolution of the difficulties experienced by these individuals in childhood. These individuals thus likely still require additional support in higher education in order to attain a level of academic success that is on par with that of their unaffected peers.

4.6.2 *Aim 2: Determination of which measures were most sensitive for discriminating dyslexic from non-dyslexic participants*

Analyses comparing the current sample of individuals with a diagnosis of dyslexia to those without yielded results that were in good agreement with previous literature on high functioning dyslexics. As expected, impairments were demonstrated on measures of phonological skills (non-word reading and non-word repetition), rapid naming and working memory. Interestingly, impairments were also demonstrated on measures of oral language skills, in particular the NEPSY oromotor sequences sub-test. This test of oromotor control requires participants to repeat short phrases and tongue twisters. Impairments in oral language were also reported in a sample of dyslexic college students by a recent study by Wiseheart and Altmann (2018) using a different kind of task; this

required participants to generate a spoken sentence using a given verb and two nouns. Dyslexic participants were found to demonstrate poorer performance in terms of slower responses and production of sentences that were less fluent, grammatical and complete.

This task used by Wiseheart and Altmann (2018) is arguably more similar to the TOAL sentence assembly task employed in the current study, which required participants to generate a spoken sentence to combine a set of shorter sentences. The current study in fact found no difference between dyslexic and non-dyslexic participants for TOAL sentence assembly. Crucially however, this task was not timed and participants had access to the written stimuli throughout construction and reporting of their response; conversely the task used by Wiseheart and Altmann (2018) assessed speed of response and stimuli were removed at the initiation of a spoken response. Working memory demands between the tasks are therefore likely to differ, which may explain the discrepant results. Indeed, when an offline version of the spoken sentence production task was used by Altmann, Lombardino and Puranik (2008), older dyslexic students were found to produce sentences that were just as grammatical as those produced by non-dyslexic students, but with poorer fluency (e.g. more pauses and reformulations). Such dysfluencies would not be captured by the TOAL sentence assembly task, since only the final full sentence produced by the participant was recorded and scored for grammaticality, regardless of how many false starts were made. The study by Wiseheart and Altmann (2018) reported that impaired performance on the online sentence production task was primarily underwritten by impairments in working memory. Importantly however, the current finding of impairments on an oromotor control task suggests that the oral

language difficulties of dyslexic adults may not necessarily be solely attributable to working memory difficulties, since this task has relatively low working memory demands.

The current findings of areas of spared functioning in my high functioning dyslexic adults are also mostly in agreement with previous literature. The robust findings of spared non-verbal IQ and vocabulary skills in dyslexic university students were replicated (Hatcher et al., 2002). Spared performance was also found for a measure of comprehension from the Expression, Reception and Recall of Narrative Instrument (ERNNI). This tests comprehension of material from within a narrative previously constructed by the participant from a picture storybook. Previous research has reported that comprehension from reading can be affected in adult dyslexics, particularly when under timed conditions (Pedersen, Fusaroli, Lauridsen, & Parrila, 2016). The current finding suggests that general comprehension of material accessed in a manner that does not involve reading is however intact. A lack of group differences were also found on more general measures of oral language, such as mean length of utterance from ERRNI storytelling and the language structure composite score of the Communication Checklist Self Report (CC-SR). This suggests that everyday language within high functioning dyslexic adults is unaffected; oral language difficulties may only manifest when put under conditions of pressure, as in the NEPSY oromotor sequences test. The previously discussed finding of normal performance on the TOAL sentence assembly sub-test is somewhat inconsistent with previous research reporting difficulties with syntactic processing in adult dyslexics (Wiseheart, Altmann, Park, & Lombardino, 2009), however again is likely due to the reduced working memory demands in this task. As expected, social and

pragmatic skills as assessed on self-report measures (CC-SR) were also unaffected in dyslexic participants.

Despite finding significant group differences across multiple areas, logistic regression analyses indicated that phonological skills as assessed by non-word reading dominated in predicting dyslexia diagnosis state in the current sample, with the addition of other measures failing to improve prediction. This fits with the prevailing view of a core phonological deficit in dyslexia, in which the pattern of impairments seen is proposed to stem from a central difficulty with phonological processing (Snowling, 2005). Previous studies of adult dyslexic students however using similar analyses have reported that improved discrimination is achieved with the addition of other variables not included in the current study. Hatcher, Snowling and Griffiths (2002) used discriminant function analysis and found that the best classification was achieved using three tests of non-word reading, writing speed and short term memory. A study by Tops, Callens, Lammertyn, van Hees and Brysbaert (2012) aimed to replicate this study in a sample of Dutch dyslexic students; using logistic regression analyses they found that the best fitting model included word reading, word spelling and phonological awareness. The absence of measures of phonological awareness, spelling and writing in the current study may explain the failure to find any benefits of additional variables in model fit compared to non-word reading alone.

4.6.3 ***Aim 3: Correspondence between diagnostic labels and data-driven sub-groups of participants***

There has been recent interest in the application of unsupervised machine learning approaches to the study of developmental disorders (see Hong, 2018). Rather than starting with *a priori* categories based on diagnostic criteria, these

data-driven approaches allow one to derive similarity structure from data to consider how cases naturally ‘cluster’ together. These have been used in a small number of studies of children with various developmental disorders (Astle et al., 2019; Bathelt, Holmes, et al., 2018), generally with the conclusion that diagnostic labels are not good predictors of data-driven cluster membership.

The current analysis adds to this literature, by demonstrating that the same conclusion is suggested when applied to an adult sample, albeit a high functioning one. The application of such a data-driven clustering analysis to data on the full sample using a select number of variables with relevance to dyslexia yielded two clusters; one that was broadly dominated by dyslexic individuals, and a second that, although including the vast majority of control participants, also contained an array from the other diagnostic categories, including all participants with autism and SpLD as well as a number of dyslexics (19 participants). Thus, although it is tempting to label the first cluster as ‘dyslexic’ and the second as ‘non-dyslexic’, one can question why such a large number of dyslexic participants were placed in the ‘control’ cluster. One possibility is that this was due to the measures chosen for the analysis; these dyslexic participants may demonstrate a slightly different profile of impairment that the current selected measures were not sensitive to, for example being more impaired on spelling, writing or phonological awareness. The addition of these variables in this type of analysis may thus result in a neater separation of dyslexic and non-dyslexic participants. Alternatively it is possible that these dyslexic adults have compensated or recovered from their difficulties to such an extent that they appear more similar to control individuals on these measures.

Overall however, this analysis highlights that not all individuals with the same diagnostic label demonstrate the same profile of impairment. In fact, the 50:50 split of dyspraxic cases between the two clusters (5 participants in each) further suggests that some of these individuals labelled as dyspraxic appear more similar to cases of dyslexia than to controls on these measures of word reading, phonological skills, working memory and oral language. This is consistent with previous research reporting difficulties in phonological processing in dyspraxic children (Dewey et al., 2002). Overall, this analysis highlights how, at least for the current unusual high functioning sample, the diagnostic label of an individual does not necessarily predict well the type of support an individual needs; it is therefore likely that support given should be assessed relative to an individual's profile of strengths and weaknesses, rather than in relation to a diagnostic label.

4.6.4 *Limitations of the study*

There are a number of limitations of the current study which constrain the interpretation of the findings. As already raised, the restricted range of measures more specifically relevant to dyslexia may have affected the results of the logistic regression analysis. In particular, measures of phonological awareness, spelling and writing would likely have yielded significant effects in predicting dyslexia, beyond the measure of non-word reading that dominated the prediction model in the current analysis. Their omission reflects the fact that this test battery was not designed to be specifically tailored to dyslexia, but rather to address a broad range of language functions that may be differentially affected in different disorders. Selection was thus constrained by the need to compile a comprehensive battery that would not be too lengthy to administer.

A further limitation of the study is the small number of participants in diagnosis categories other than dyslexia, with only ten cases of pure dyspraxia, four of Autism and two of Specific Learning Difficulty. The sample was recruited predominantly through university disability advisory services with the aim of collecting a sample with a broad range of language, literacy and communication skills rather than with regard to specific diagnoses. The overrepresentation of dyslexia in the resulting sample was thus unexpected and not reflective of skewed recruitment sources. This limits interpretation of some of the findings, particularly those from the data-driven clustering analysis; a repeat of this analysis with a larger more balanced sample of individuals with dyslexia, dyspraxia and no diagnosis would be of interest to more comprehensively investigate relationships between these two interrelated disorders.

It should be noted that this study differs from similar literature in that the comparison group for the analyses focusing on dyslexia contained participants both with no diagnosis and with diagnoses of other disorders. This may have dampened differences between groups, given frequent sharing of symptoms between these disorders. Indeed, some effect sizes reported in the current study are lower than those that have been previously reported when comparing with control (non-disordered) samples (see Swanson & Hsieh, 2009). Participants with diagnoses of other disorders were included so as to maximise power, and with the rationale of identifying those measures that were specific to a diagnosis of dyslexia, rather than being common to other disorders. Overall, when considering the power of this study, it should be noted that no power calculations were done *a priori*, since this was a secondary analysis of existing data collected for a

different study (see *Chapter 5*). Therefore, all analyses should be viewed as exploratory and results interpreted accordingly.

Lastly, it should be acknowledged that this sample of adults with developmental disorders is likely a highly skewed sample. Recruitment was predominantly via the Disability Services at the University of Oxford, resulting in a sample with a large proportion of high achieving adults with developmental disorders. This may explain why a relatively modest proportion of the DD group demonstrated impaired language functioning relative to standardised population norms. It should be further pointed out that adults with a history of learning disabilities who reach higher education are in general likely to be a highly selective sample compared to the general population with learning difficulties. For example, a recent study by Chatzitheochari and Platt (2019) reported that the level of parental education in a sample of adolescents with disabilities who had continued in full-time upper secondary education was in fact higher than that of their counterparts without disabilities, suggesting a skew towards high SES among these individuals. This means that the findings from the current analyses should not be interpreted as representative of the wider population of adults with developmental disorders; instead these should be viewed as providing further characterisation of the nature of strengths and weaknesses shown by the current unusual sample, which can be used to inform interpretation of subsequent chapters in this thesis that study these individuals.

This raises the general point that factors other than profiles of abilities are likely to influence educational attainment in young adults with developmental disorders. One can make a broad distinction between primary effects, in which lower ability directly constrains academic achievement, and secondary effects, such as the

educational choices these students make from the range of choices available to them given their level of ability. Chatzitheochari and Platt's (2019) longitudinal study of adolescents with disabilities (including a large number with dyslexia) suggested that both primary and secondary effects influence educational outcomes. Whilst prior attainment at Key Stage 2 was the primary predictor of success in attaining GCSEs, expectations of academic success was found to be the main predictor of whether students chose to continue to A-levels after GCSEs. The educational success of students with learning disabilities can thus be considered as constrained both by primary effects of reduced abilities but also secondary effects of expectations and stigma that affect the educational choices they make. Future research will need to consider both primary and secondary effects when investigating the factors that predict success in entering and completing higher education in samples with developmental disorders. The assumption that the current high achieving sample will demonstrate a less severe pattern of impairment than adults with the same disorders who do not successfully enter higher education ignores the potential role of these other secondary factors in determining academic success.

4.6.5 *Summary and conclusions*

The findings presented here provide insight into the profile of impairments demonstrated by high functioning adults with a history of developmental disorders. This adds to the growing body of evidence that despite the fact that these individuals can compensate for their difficulties to a point that enables them to achieve academic success in higher education, they still experience areas of weakness that are likely to put them at a disadvantage relative to their fellow students. Difficulties with reading familiar and unfamiliar words as well as

limitations in working memory in students with dyslexia would be expected to both hinder learning from written material (inherent to most courses of study) and result in poorer performance under timed exam conditions. This suggests that the common practice of allowing dyslexic participants extra time in exams is justified. However, the preliminary evidence presented in the current data-driven clustering analysis with a small sample of dyspraxic adults further suggests that some of these individuals can demonstrate a similar profile of impairment on these measures, whilst other individuals with a dyslexia diagnosis may be more similar to control individuals. This highlights that the type of support given to an adult student with a developmental disorder should be based on their individual needs, rather than their diagnostic label. More research is needed to systematically compare profiles of impairment between dyslexic and dyspraxic adults, to consider the shared and distinct difficulties faced by these individuals and the implications for study support. Overall, when considering the pattern of spared and impaired abilities present in 'high functioning' learning disabled samples, it is important to consider that factors other than level of ability may affect academic outcomes, such as stigma, expectations and SES.

Chapter 5 Inconsistency in lateralisation of language functions: a risk factor for language impairment?

Abstract

Disruption to language lateralisation has been proposed as a cause of developmental language impairments, however this has so far only been investigated by looking for laterality strength differences between groups on a single task. This study aimed to test the idea that poorer language abilities are associated with inconsistent lateralisation across different language functions. 104 participants (67 with a developmental disorder affecting language and 37 controls) were tested with a language assessment battery and Doppler ultrasound measurement of laterality for three language tasks that engage different language sub-processes (sentence generation, phonological decision and semantic decision). Dividing the whole sample into those with consistent versus inconsistent lateralisation across the three tasks did not find evidence that an inconsistent pattern of lateralisation was associated with poorer language skills as assessed by the chosen language measures. However, the developmental disorder group demonstrated an increased incidence of inconsistent lateralisation and evidence of weaker correlations between laterality indices across pairs of tasks. Strength of lateralisation did not differ between developmental disorder and control groups. The findings suggest that having inconsistent lateralisation for different language sub-processes does not increase risk for impaired language; however, experience of a developmental disorder may

increase the likelihood of having an unusual, more distributed brain organisation for language, potentially reflective of compensatory reorganisation.

5.1 Introduction

There has been much interest in the possible relevance of individual variability in lateralisation for language functioning. The approach of most imaging studies in this area has been to compare the language abilities of individuals divided into 'typical' and 'atypical' laterality groups on the basis of whether they show the expected left-sided pattern of brain activity during performance of a single language task. Similarly, studies of language disorders typically compare the incidence of atypical laterality in a language impaired group to that in a control group on a single lateralising task. Both approaches have generally produced inconsistent findings across studies. This evidence has been reviewed in detail in the Introduction of this thesis (see *Introduction: 1.3.1 Is lateralisation beneficial?*); the key aspects of this literature are summarised here for the reader's convenience.

5.1.1 *Defining 'atypical' laterality*

'Atypical laterality' is a generic term that is used to describe individuals showing either a right-bias or a lack of hemispheric dominance (bilaterality) on such standard laterality measurements. However, individuals may differ in the organisation of their language systems across the hemispheres in meaningful ways that are not captured by this simple typical-atypical dichotomy. Work on anatomical asymmetries by Ide et al., (1999) observed that individual differences in asymmetries of fissurization patterns across different brain regions were not associated with one another. They concluded that brain laterality may thus reflect the summation of multiple local asymmetries that develop independently, rather

than at a hemispheric level. More recent fMRI work has further reported cases of variability in the extent and side of laterality measured across different language regions and different language tasks within an individual (see *Introduction: 1.5 Variability in lateralisation of different language sub-processes* for a review of this literature). Tailby et al., (2017) argued that in light of such cases of inconsistent laterality, trying to determine "which is the dominant hemisphere?" for a given individual is an ill-posed question that incorrectly assumes uniformity in laterality throughout the language network. This work suggests the possibility that different aspects of language can lateralise independently, requiring revision of the construal of lateralisation as a unitary construct (Bradshaw, Bishop, et al., 2017; Bradshaw, Thompson, et al., 2017; Woodhead, Bradshaw, Wilson, Thompson, & Bishop, 2019).

5.1.2 *Language lateralisation and language impairment*

Disruption to language lateralisation has been proposed as a cause of developmental language impairments by multiple theories over the last century (Annett, 1985; Orton, 1937). Indeed, recent research has reported a higher incidence of atypical laterality in samples of language impaired individuals (Bishop, Holt, Whitehouse, & Groen, 2014; de Guibert et al., 2011; Illingworth & Bishop, 2009; Sato et al., 2011). However, overall this evidence has been mixed; in particular, a recent large-scale study of children with developmental language disorder by Wilson and Bishop (2018) failed to find evidence of reduced laterality compared to typically developing children using fTCD with an animation description task. Here, I consider whether consistency rather than degree of language lateralisation may be important, given evidence that different components of language processing can lateralise independently.

A theory suggesting that variability in lateralisation for different language processes may constitute a risk factor for impaired language was proposed by Bishop *et al.*, (2014). This 'Left Brain Bias model' proposes that different language processes develop lateralisation independently via a probabilistic process. The majority of individuals are proposed to have a 'left brain bias' phenotype, in which the probability of development of laterality for each process is weighted towards the left hemisphere. Conversely, a minority of individuals are proposed to have a 'no bias' phenotype, in which each language process has equal likelihood of developing left or right lateralisation. For such individuals therefore, there will be increased variability in the side of lateralisation of different language functions. This could represent a less efficient language network, increasing the need for hemispheric integration and transfer of information across the corpus callosum.

Indeed, many theories have suggested that there may be evolutionary pressures promoting common lateralisation of subsystems within a domain for optimal functioning. Theories taking a neuroanatomical approach have suggested that lateralized neural networks may have evolved in response to increases in callosal transmission delays associated with larger brains (Aboitiz et al., 2003; Ringo et al., 1994). This may have created an evolutionary pressure for functions that rely on multiple frequent interactions to be contained within the same hemisphere, so as to depend on relatively faster within-hemispheric tracts. Similarly, Kosslyn (1987) proposed a computational theory for understanding hemispheric specialisation and argued that an innately specified leftward bias for a speech control centre leads to a snowball effect in which downstream subsystems that interact with this centre will tend to develop congruent lateralisation. That is, the need for efficient communication between subsystems that serve language will

create a pressure towards a common lateralisation pattern across functions. Thus, a high level of variability in lateralisation of different language processes would be expected to result in reduced efficiency of language processing, leading to impaired language function.

The study of consistency in lateralisation within individuals can take two approaches. Firstly, studies have used a single language task with fMRI to look at dissociations in laterality at a regional level, for example comparing frontal with posterior lateralisation for the same language task (Berl et al., 2014; Bethmann et al., 2007). Secondly, inconsistency in lateralisation can be addressed at the language processes level, by comparing laterality across multiple language tasks (Badcock, Nye, & Bishop, 2012; Häberling et al., 2016; Stroobant, Buijs, & Vingerhoets, 2009). It appears that neither approach has been taken to the study of the relationship between laterality and language function in a language impaired sample. A study by de Guibert et al., (2011) did measure laterality in a group of children with specific language impairment using a panel of four language tasks; however, these were simply used in a combined task analysis, without any comparison of individual patterns of laterality across the tasks.

5.1.3 *Aims of the current study*

Deciphering the nature of the disruption to lateralisation in language impairment may require the consideration of differences in laterality patterns within individuals, either across different regions of the language network, or across different tasks engaging distinct language processes. This study takes this second approach by using functional transcranial Doppler sonography (fTCD) to investigate individual differences in the pattern of lateralisation for different language functions, and how these may relate to language ability. Lateralisation

was measured within the same individuals across different language tasks that engaged distinct components of language functioning, whilst being closely matched on non-linguistic parameters. These laterality measurements were used to compare the language abilities of those with consistent dominance across tasks with those of individuals with inconsistent dominance, within a sample containing an overrepresentation of language impaired individuals. It was hypothesised that individuals with inconsistent language lateralisation would show poorer language skills than those with consistent lateralisation.

5.2 Methods

This study was pre-registered prior to data analysis on Open Science Framework (https://osf.io/fvhxq/register/565fb3678c5e4a66b5582f67). Deviations from this pre-registration are described in the relevant sections of text and summarised at the end of the methods section.

5.2.1 *Materials and procedure*

Participants underwent two testing sessions; a first session in which a battery of language assessments was administered, and a second session in which functional transcranial Doppler sonography (fTCD) was used for laterality measurement. These sessions were either administered on the same day, or with a gap ranging up to 16 months apart.

5.2.2 *Measurement of language skills*

The assessments used to measure language, literacy and memory skills have already been described in a previous chapter (*Chapter 4: 4.2.3 Language assessment battery*). The 12 measures included in the current analysis were highlighted in bold in Table 4.1 of this previous chapter. Note that one of these

measures was changed from that pre-registered; TOWRE non-word reading was replaced by TOWRE overall standard score due to ceiling effects in the former. It has been demonstrated in Chapter four that these 12 measures were sensitive in discriminating between the developmental disorder and control groups reported on in the current chapter.

In addition to these language measures, a quantitative measure of handedness was obtained using the Annett peg moving task (Annett, 1985), supplemented by the Quantification of Hand Preference (QHP) task (Bishop et al., 1996) and the Edinburgh handedness inventory (Oldfield, 1971). Tests one and two of the Cattel Culture fair test (Cattell & Cattell, 1973) were also administered in order to obtain a measure of non-verbal reasoning.

5.2.3 *Measurement of lateralisation*

Lateralisation of language function was measured using functional transcranial Doppler ultrasound (fTCD). FTCD is a reliable method of laterality measurement which uses ultrasound waves to insonate the middle cerebral artery (MCA) of each hemisphere in order to measure changes in the speed of blood flow in right and left hemispheres (Deppe, Knecht, Henningsen, & Ringelstein, 1997; Knecht et al., 1996).

Tasks were selected from a larger battery of tasks for fTCD laterality measurement reported on by Woodhead, Bradshaw, Wilson, Thompson and Bishop (2019). This study measured test-retest reliability for 6 different language tasks in typically developing individuals, and used a structural equation modelling analysis to explore the factor structure of task LIs. This allowed selection of tasks for the current study on the basis of the following criteria: (1) demonstration of

reliable laterality measurement, (2) engagement of different key aspects of language functioning whilst being closely matched on non-linguistic parameters and (3) a tendency to lateralise together in typically developing individuals. These guidelines were used to select the following three tasks; sentence generation, semantic decision and phonological decision. These tasks were found by Woodhead et al., (2019) to yield good levels of test-retest reliability (ranging from 0.57 to 0.84), and were designed to engage distinct aspects of language processing; namely, speech production, phonological processing and semantic processing.

The design of these tasks was guided by the findings of the systematic review of fMRI laterality tasks presented in Chapter three of this thesis. They were designed so as to minimize differences in non-linguistic parameters and thus engagement of non-linguistic processes that may contribute to differences in the measured LIs. All tasks were cued by similar black line-drawing stimuli, in order to closely match perceptual processing across tasks. In particular, the two decision tasks used highly similar sets of picture stimuli representing words that were matched across task sets on the following psycholinguistic variables using N-watch: familiarity, orthographic neighbourhood, imageability, number of phonemes and frequency. All three tasks were also designed so as to follow a similar structure (illustrated in Fig. 5.1). Each task was made up of 15 trials, each 33 seconds long. Trials started with the word ‘CLEAR’ on screen for 3 seconds, instructing participants to clear their mind in preparation for the next trial. This was followed by a 20 second period during which the task is performed. The procedures for each task type are detailed below. After the task period for that

trial had finished, the word ‘REST’ appeared on screen for 10 seconds. Participants were instructed to try and keep a clear mind during this rest period.

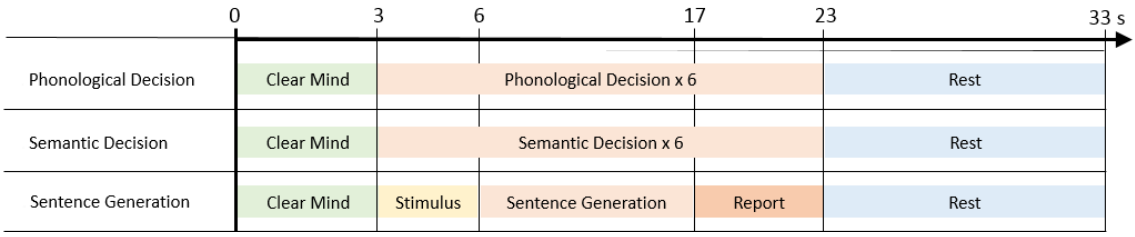


Figure 5.1: FTCD tasks. *Task design and timings of the three language tasks for use with fTCD.*

5.2.3.1 Phonological decision

The phonological decision (PD) task was based on a rhyming decision task used by Bishop and Robson (1989). This required participants to make a rhyme judgement on a pair of words represented by pictures. The pictures consisted of line drawings of single syllable words that were designed to be easily nameable, arranged into 45 rhyming pairs and 45 non-rhyming pairs. These pictures were mostly taken from the International Picture Naming Project (IPNP) database (Szekely et al., 2004). An example of a picture pair stimulus used in this task is shown in Figure 5.2A. Rhyming and non-rhyming pairs were designed to not differ significantly in degree of orthographic overlap, using the Davis (2000) similarity metric (see http://www.pc.rhul.ac.uk/staff/c.davis/Utilities/MatchCalc/). For each trial, a series of 6 of these picture pairs was presented sequentially, each for 3.33 seconds (totalling 20 seconds). For each pair, the participant had to decide whether the words represented by the pictures rhymed or not. Participants provided their decision by responding with a ‘Y’ key if the stimuli rhymed or an ‘N’ key if the stimuli did not rhyme. Participants were required to respond within the 3.33 second window of stimulus presentation.

5.2.3.2 Semantic decision

The semantic decision (SD) task required participants to make a semantic category judgement about objects represented in a pair of pictures. This task was designed to closely match the phonological decision task. Six picture pairs of the same format as in the phonological decision task were presented, again each for 3.33 seconds. For each pair, the participant decided whether the objects depicted were from the same semantic category or not (e.g. both types of food). Participants were instructed to respond with a 'Y' key if the stimuli came from the same semantic category or an 'N' key if the stimuli belonged to different categories. An example stimulus for the semantic decision task is shown in Figure 5.2B.

5.2.3.3 Sentence generation

The sentence generation (SG) task was based on an fMRI paradigm developed by Mazoyer *et al.*, (2016) which required participants to generate a spoken sentence to describe a presented line drawing. The task began with presentation of a simple black line drawing on a grey background for 3 seconds. An example of one such line-drawing stimulus is depicted in Figure 5.2C. This was followed by presentation of a black cross in the centre of the screen for a duration of 11 seconds, during which time the participant was required to covertly generate a simple sentence to describe the picture. Participants were trained in advance to generate sentences of a particular structure: beginning with a subject (e.g. "the boy"), followed by a description of the subject (e.g. "with marbles"), a verb (e.g. "plays") and ending with a detail about the action (e.g. "on the floor"). Following this covert generation, the word "REPORT" was presented on screen, prompting participants to overtly report their sentence during a period of 6 seconds.

For the experiment, participants were first set up with the fTCD headset and probes, before they performed each of the language tasks in turn. The order of performance of these tasks was randomised across subjects. All stimuli were presented using Psychtoolbox (Brainard, 1997) for Matlab 2012a (Mathworks Inc, Natick, MA). Each task lasted around 15 minutes, giving a total session time of 1 hour to 1 and a half hours depending on the length of time required for fTCD set-up and identification of the MCA.

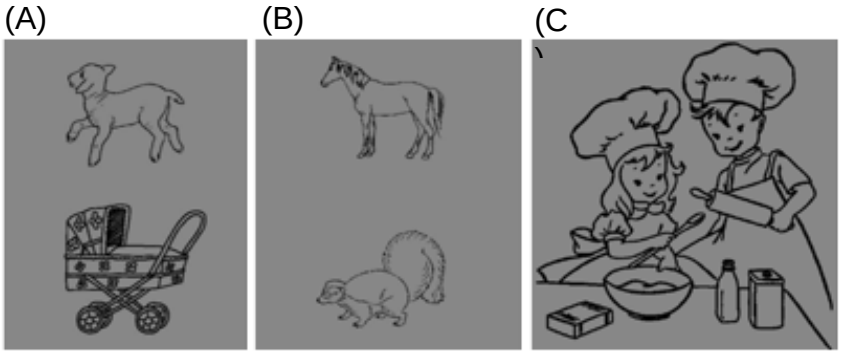


Figure 5.2: FTCD language task stimuli. *Example picture stimuli used in the (A) phonological decision, (B) semantic decision and (C) sentence generation tasks. The majority of the picture stimuli were taken with permission from the International Picture Naming Project (Szekely et al., 2004).*

5.2.3.4 FTCD analysis and calculation of laterality indices

The methods described here for calculation of fTCD laterality indices are the same as those employed in the study by Woodhead et al., (2019). For the purposes of analysis of the fTCD data, two time periods were defined within each epoch: a baseline period from -6 to -1 seconds (used for baseline normalisation), and a period of interest (used for LI calculation) which varied depending on the task due to differences in the stimulus timings; for sentence generation this was from 6 to 20 seconds (ending when the 'REPORT' stimulus was presented), and

for the decision tasks this was from 6 to 26 seconds. Time 0 seconds represents the onset of the 'CLEAR' stimulus at the start of each trial.

An R script was used to analyse this data and calculate an LI value for each participant on each of the three tasks individually (this can be found at: https://osf.io/k7nhz/). This script starts by downsampling the data to 25 Hz by taking every 4th data point. The script identifies periods of drop-out or spiking within each epoch using lower and upper limits set to the .0001 and .9999 quantiles of the mean for that epoch. A single extreme value identified within an epoch is replaced with the mean signal value; if more than one extreme value is identified the whole epoch is excluded. In order to account for the possibility of one side's signal being stronger than the other due to differences in the angle of insonation of the left and right middle cerebral arteries, the data within each channel is normalised by multiplying by 100 and dividing by the overall data mean; this causes each channel to have a mean of 100 whilst preserving its variance. A heartbeat correction is applied that extracts variation in the signal due to the cardiac cycle, followed by a baseline correction achieved by subtracting the mean for the baseline period from the data-points for each channel independently. The script then checks that all data-points are within 60% and 140% of the mean, and excludes any epochs that contain data-points outside of this range. An average is then taken across epochs for each data-point to give an average time-course for left and right channels. The difference between channels (L-R) is then taken at each data point. Figure 5.3 shows the grand mean L-R difference for each task, using data averaged across all trials and all subjects.

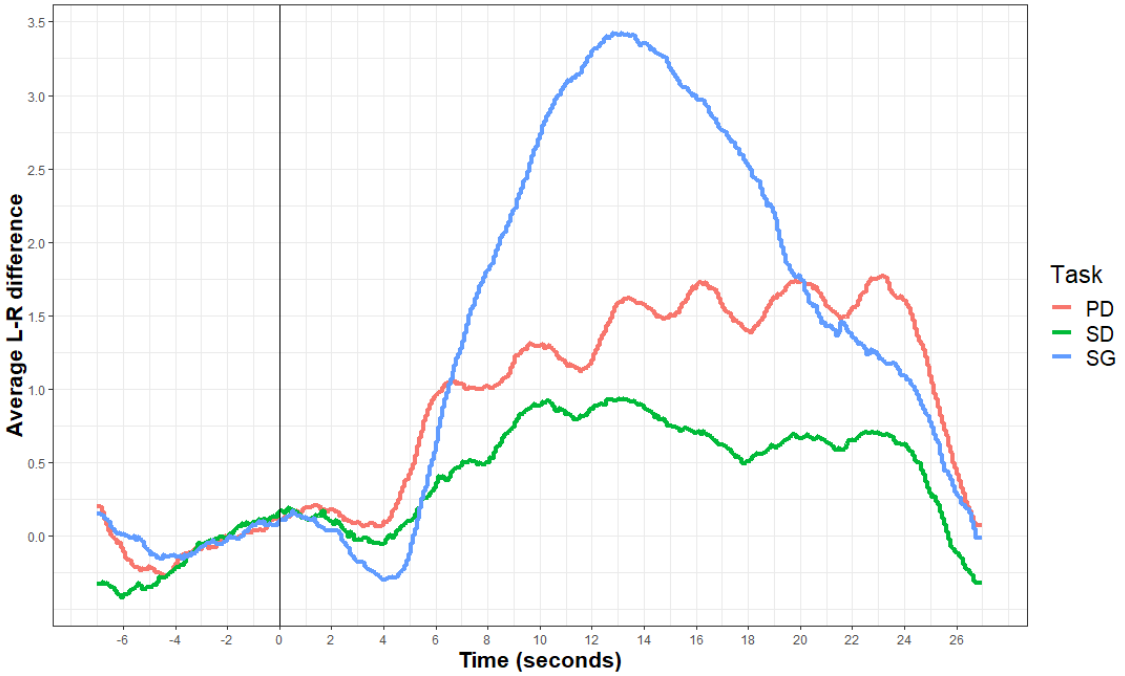


Figure 5.3: Average left minus right difference curves. *Plot shows mean difference between the signal in the left and right channels averaged across trials and participants. Time 0 represents the onset of the 'clear mind' instruction.*

Each individual's L-R average difference curve was then used to calculate a laterality index for each task. The original pre-registered analysis planned to use the 'peak' method, in which the point of maximum difference between L and R channels within the period of interested is identified. A time window of 2 seconds centred on this peak difference is then taken. The mean of the L-R differences within this time window is then used to give the laterality index for that participant. A different method of LI calculation known as the 'mean' method takes an average of this difference across the whole period of interest; this provides a measure which is more representative of the overall trend for laterality within the period of interest. The other key benefit of this mean LI measure is that it yields LI values which are normally distributed, whereas the peak method LIs show a bimodal distribution (see *5.3.2 LI values* for a demonstration of this). LIs were

calculated using both methods, but mean LIs were used in the analyses due to these advantages.

The calculated LI values were used to classify participants into two dominance categories for each task, left lateralised (L) or right lateralised (R), according to whether their LI value was positive or negative respectively. It was decided not to include bilateral as a dominance category due to concerns as to the reliability of this classification (see Jansen *et al.*, 2006 and Bradshaw, Bishop, *et al.*, 2017, for discussion of this in the fMRI literature). In the data reported by Woodhead *et al.,* (2019) average concordance for dominance across two testing sessions measured using these three tasks was 88% when only left and right categories were used.

Each participant was then classified as either consistently lateralised or inconsistently lateralised according to the concordance in their dominance classifications across the three tasks. A participant was classed as 'consistent' if all three of their dominance classifications for the different tasks agreed. A participant was classed as 'inconsistent' if there was discordance in their dominance classification for the three tasks.

5.2.4 ***Analysis plan***

All analyses detailed here were conducted using the statistical software R (R Core Team, 2019). An R Markdown script for running all analyses presented in this chapter can be found on Open Science Framework (https://osf.io/k7nhz/).

5.2.4.1 Pre-registered analyses

The main analysis of this study tested the hypothesis that variability in the side of lateralisation of different aspects of language is associated with poorer language

skills. The whole sample (composed of both language impaired and non-impaired individuals, see *5.2.5.1 Details of Recruited Sample* for more details) was divided into those with consistent and inconsistent laterality, as outlined above. The original analysis plan was to compare consistent and inconsistent lateralisation on the 12 language measures using the multivariate counterpart of the t-test, the Hotelling's T-square test (see Schumacke, 2016). Hotelling's T-square gives an index of the difference between two groups as assessed on a set of correlated variables, while controlling the type I error rate and taking the correlation between variables into account. A significant Hotelling's T-square statistic would indicate that the two groups were significantly different on the language measures as a set.

However, after examination of the collected data on the language measures, it was found that these violated the assumption of normality, with cases of skew and bimodality (see *Chapter 4: 4.2.4 Normality of the measures*). Use of this parametric test on this non-normal data would therefore result in poor estimation of standard errors and thus a biased test statistic estimate, inflating the likelihood of type 1 errors. While transformations such as the Box-Cox transform are able to deal with skewness, they are not successful at inducing normality in bimodally distributed measures. Therefore, instead a non-parametric multivariate test was used to compare the language measures between laterality groups (Oja & Randles, 2004; Sirkiä & Taskinen, 2007, see *Chapter 4: 4.3.1 Statistical analyses* for full description). This non-parametric multivariate test was run using the R package SpatialNP (Sirkiä et al., 2007) to compare performance of the two laterality groups on the 12 language measures.

To investigate the potential relationship between handedness and consistency of language lateralisation, strength of right handedness was compared between consistent and inconsistent groups measured using the Annett peg-moving task. This uses speed of performance with the right and left hand (to move wooden pegs from one row of a peg board to the other) to calculate a handedness quotient of ((L-R)/(L+R)*100). More positive/negative numbers indicate faster performance (higher skill) with the right/left hand respectively. This normally distributed measure of relative hand skill was compared between the consistent and inconsistent groups by means of an independent t-test, with the prediction that the inconsistent group will demonstrate lower relative right hand skill than the consistent group.

If a significant group difference between consistent and inconsistent groups was found on the language measures, two additional analyses were planned to rule out the effects of potential confounding factors. Firstly, the groups would be compared on their scores of non-verbal IQ as measured by the Cattell Culture Fair test, to test the possibility that group differences in ability are not specific to language. Secondly, to check for the effects of task performance on laterality measurement, Spearman's correlations would be run between LI and performance measures on these tasks, with the hypothesis that there would be no correlation (using a Bonferroni correct p value of .05/3).

5.2.4.2 Non-preregistered exploratory analyses

Further analyses that were not originally pre-registered were carried out on the data. In order to explore relationships between task LIs, Pearson's correlations were run for pairs of task LIs in the developmental disorder and control groups separately. This was used to test the hypothesis that correlations between task

LIs were weaker in the developmental disorder than the control group. Further, to investigate whether the developmental disorders group demonstrated a significantly increased frequency of inconsistent laterality patterns, a chi-squared test was used to compare frequencies in the inconsistent and consistent laterality groups from the two diagnosis groups.

One possibility is that individuals with developmental disorders may not only be inconsistent in which hemisphere they use across different language tasks, but also across time within performance of a single task. To investigate this, the percentage of trials within each task that were left lateralised were calculated for each individual. A metric of consistency in lateralisation over time was then obtained by taking the absolute difference between this value and 50. This gave a value from 0 to 50, where 0 indicates maximal inconsistency (50/50 split in terms of which side was used across trials), and 50 indicates complete consistency (one side was used on 100% of trials). Mean values were then compared between control and developmental disorder groups for each task using independent t-tests, with the exploratory hypothesis of a reduced level of consistency over time in the developmental disorder group.

Lastly, although the main hypothesis of this study focused on the relationship between language abilities and consistency in lateralisation, the data were further analysed in a more conventional way to test for a relationship with strength of lateralisation. A multi-level modelling analysis was used to compare strength of laterality between the developmental disorder and control groups across the three laterality tasks.

5.2.5 ***Participants and statistical power analysis***

The pre-registered power analysis to determine sample size was based on the use of the Hotelling's T-square test, as originally planned. This first required an estimate of effect size. Since there was no previous literature using this kind of analysis to look at the relationship between consistency in lateralisation and language ability, a moderate effect size of d = 0.5 was chosen, which represents the size of effect which would be of theoretical interest. On the basis of previous work using these tasks (Woodhead et al., 2019) it was predicted that the numbers of participants falling into consistent versus inconsistent laterality groups would follow a ratio of 2:1; that is, it was expected that around two thirds of the sample would show consistent lateralisation whereas one third would show inconsistent lateralisation across the tasks. Based on this effect size and ratio of participants in the two groups, a power level of 0.9 using 12 language measures yielded a required sample size of 65 for the consistent laterality group and 33 for the inconsistent laterality group.

5.2.5.1 Details of recruited sample

In order to reach the required level in the inconsistent group, 109 participants were recruited and tested in this study. In four of these participants, it was not possible to identify a robust signal from the MCA due to increased thickness of the temporal bone window. A further participant was excluded due to previous infection of the cerebellum. This left a total of 104 participants. Details of this sample have already been given in a previous chapter (*Chapter 4: 4.2.1 Participants*). As previously outlined, about two thirds of this sample had a diagnosis of a developmental disorder (67 participants) and around one third individuals had no impairment (37 participants).

5.2.5.2 Summary of departures from pre-registered methods

1. **Change in one of the 12 language measures.** The pre-registered measure of TOWRE non-word reading was replaced by TOWRE overall standard score due to ceiling effects in the former.
2. **Laterality index calculation method.** The pre-registration stated that the peak method of LI calculation would be used, as is conventional in fTCD analyses. However, as outlined above, the decision was made to use a newly developed mean method of LI calculation, as this provides a more representative measure of laterality for the individual, and a more normally distributed measure at the group level.
3. **Choice of multivariate test to compare the consistent and inconsistent laterality groups on language measures.** The pre-registered analysis to compare laterality groups on the 12 language measures used the parametric Hotelling's T-square test. However, since data on the language measures were found to be non-normal, a non-parametric equivalent of this multivariate test was used for this comparison.

5.3 Results

5.3.1 *Behavioural data*

The following performance measures were taken for the fTCD tasks; accuracy (percentage correct) and reaction times for the two decision tasks, and average number of words generated across trials for SG. Means for each group across these performance measures are given in Table 5.1. Independent t-tests found no significant differences between groups on any of the measures (lowest *p* value = .255). Comparisons between the two decision tasks across the whole group did however find significant differences, with the PD task demonstrating significantly

lower accuracy ($t(103) = -7.287$, $p < .001$) and slower reaction times ($t(103) = 28.368$, $p < .001$) than the SD task. Correlation analyses however found no significant relationships between performance measures and laterality indices (calculated using the mean method), as can be seen in Table 5.2.

Table 5.1: FTCD task performance measures. *Means (and standard deviations) for performance measures across the three fTCD tasks for each group. PD = phonological decision, SD = semantic decision, SG = sentence generation.*

Group	PD accuracy (%)	PD RT	SD accuracy (%)	SD RT	SG number of words
Control	89.9 (5.34)	1.73 (0.19)	94.3 (6.16)	1.18 (0.22)	9.61 (0.81)
Developmental Disorder	88.7 (8.39)	1.71 (0.23)	93.9 (5.07)	1.23 (0.26)	9.65 (1.39)

Table 5.2: FTCD LI correlations. *Pearson's correlations between laterality indices (calculated using the mean method) and performance measures for each task. PD = phonological decision, SD = semantic decision, SG = sentence generation.*

	r	p value
PD LI and accuracy	0.087	0.379
SD LI and accuracy	0.053	0.588
SG LI and number of words	0.077	0.435

5.3.2 *LI values*

As previously described, two methods were used to calculate laterality indices from the fTCD data, the pre-registered 'peak' method and the new 'mean'

method. LI values on each of the three tasks are given in Figure 5.4 for the peak method and Figure 5.5 for the mean method. As can be seen, the peak method causes an absence of LI values around zero, resulting in a bimodal distribution. This is because even if values are generally equal across left and right channels throughout the period of interest, any period (however short-lived) at which there is increased signal in one channel over the other will have a disproportionate effect on the LI calculated. Conversely, the mean measure that considers the overall average L-R difference across the period of interest is much more normally distributed and results in an increase in values around zero. This more normally distributed 'mean' LI measure was therefore used in subsequent analyses. Peak and mean LI values were highly correlated (r = .965, p < .001), and the use of mean over peak LIs did not change the outcome of analyses; it did mean however that the data met assumptions of statistical tests.

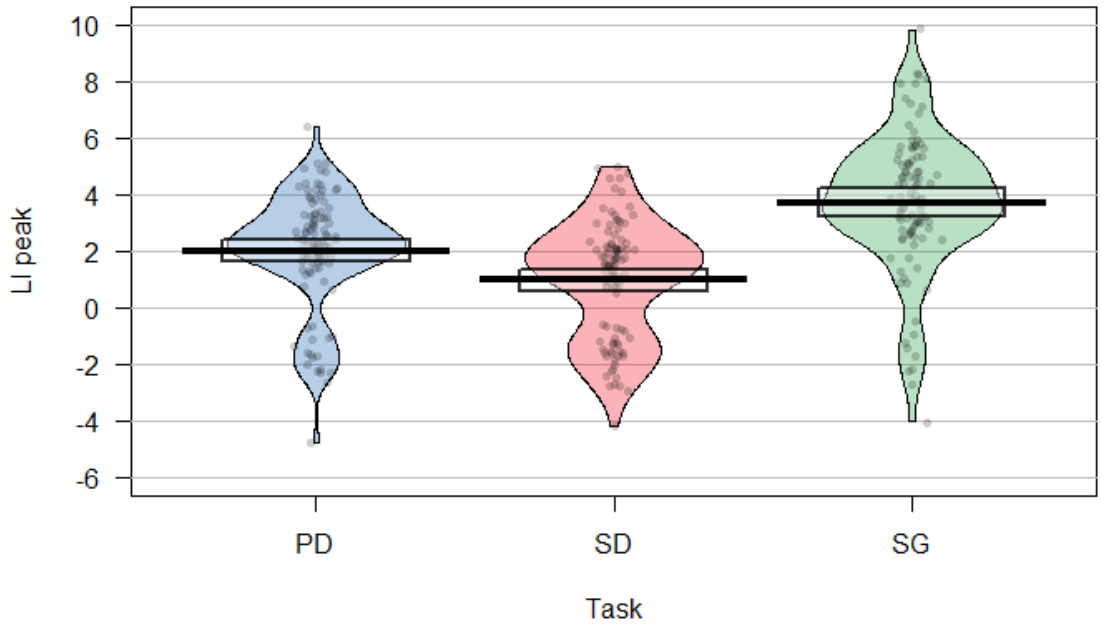


Figure 5.4: Peak method of LI calculation. *LI values across the whole sample for each task, as calculated using the peak method. Bars indicate the mean, shaded boxes indicate 95% confidence intervals. PD = phonological decision, SD = semantic decision, SG = sentence generation.*

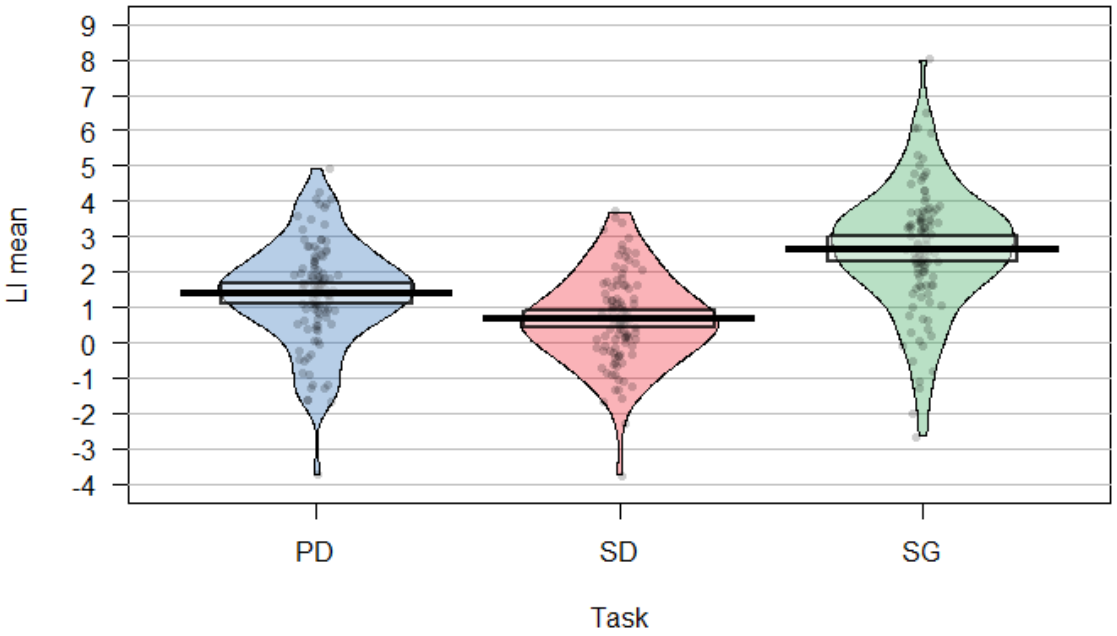


Figure 5.5: Mean method of LI calculation. *LI values across the whole sample for each task, calculated using the mean method. Bars indicate the mean, shaded boxes indicate 95% confidence intervals. PD = phonological decision, SD = semantic decision, SG = sentence generation.*

5.3.3 ***Pre-registered Analyses***

5.3.3.1 Consistency in lateralisation

Within the sample, 31 individuals were identified with inconsistent lateralisation across the three tasks, and 73 individuals with consistent lateralisation (collapsing across diagnosis groups). For a small number of participants (four individuals), classifications into consistent and inconsistent groups based on peak and mean LI values did not agree. This meant that although the required sample size of 33 in the inconsistent group was reached based on peak LI classifications, only 31 participants were classified as inconsistent using the mean measure. This small reduction was considered still acceptable given the power calculation that yielded these required sample sizes was based on a high power level of 0.9.

The most common pattern in the inconsistent laterality group was left lateralised for PD and SG and right lateralised for SD (21/31 individuals). In general, SD was more likely to dissociate from the other two tasks; in contrast PD and SG tended to lateralise together, especially if PD was left lateralised.

5.3.3.2 Comparison of language measures between consistent and inconsistent groups

Standardised (z-score) mean differences between these consistent and inconsistent laterality groups on each of the 12 language variables are given in Table 5.3. Standardised scores in each laterality group are given in Figure 5.6 as boxplots. To compare the two laterality groups on these 12 language measures, a non-parametric multivariate test was used with the raw data scores (see *5.2.4.1 Pre-registered analyses*). This found no statistically significant difference between groups across the measures (Q^2 = 6.49, *df* = 12, *p* = 0.889). This indicates that the inconsistent and consistent groups do not differ in their language abilities as assessed by the measures used in this study.

5.3.3.3 Consistency in lateralisation and handedness

To investigate whether handedness related to degree of consistency of lateralisation, an independent t-test was run to compare degree of right handedness as measured by the Annett peg moving task between consistent and inconsistent groups. This found no significant group difference in right handedness, $t(69.71) = 0.045$, $p = 0.965$.

Table 5.3: Comparison of laterality groups. *Standardised (z-score) differences between the two laterality groups (consistent versus inconsistent) on the 12 transformed language variables.*

Language measure	Standardised difference between laterality groups	Standard error
NEPSY oromotor sequences	-0.037	.215
NEPSY non-word repetition	-0.035	.215
Digit span (total score)	-0.285	.214
TOWRE overall standard score	-0.375	.212
YAR-R objects (rate)	-0.349	.213
ERRNI MLU	0.018	.215
ERRNI comprehension	-0.126	.215
TOAL sentence assembly	-0.251	.214
WASI vocabulary	-0.088	.215
CCSR language structure	-0.203	.214
CCSR pragmatic skills	-0.208	.214
CCSR social engagement	-0.237	.214

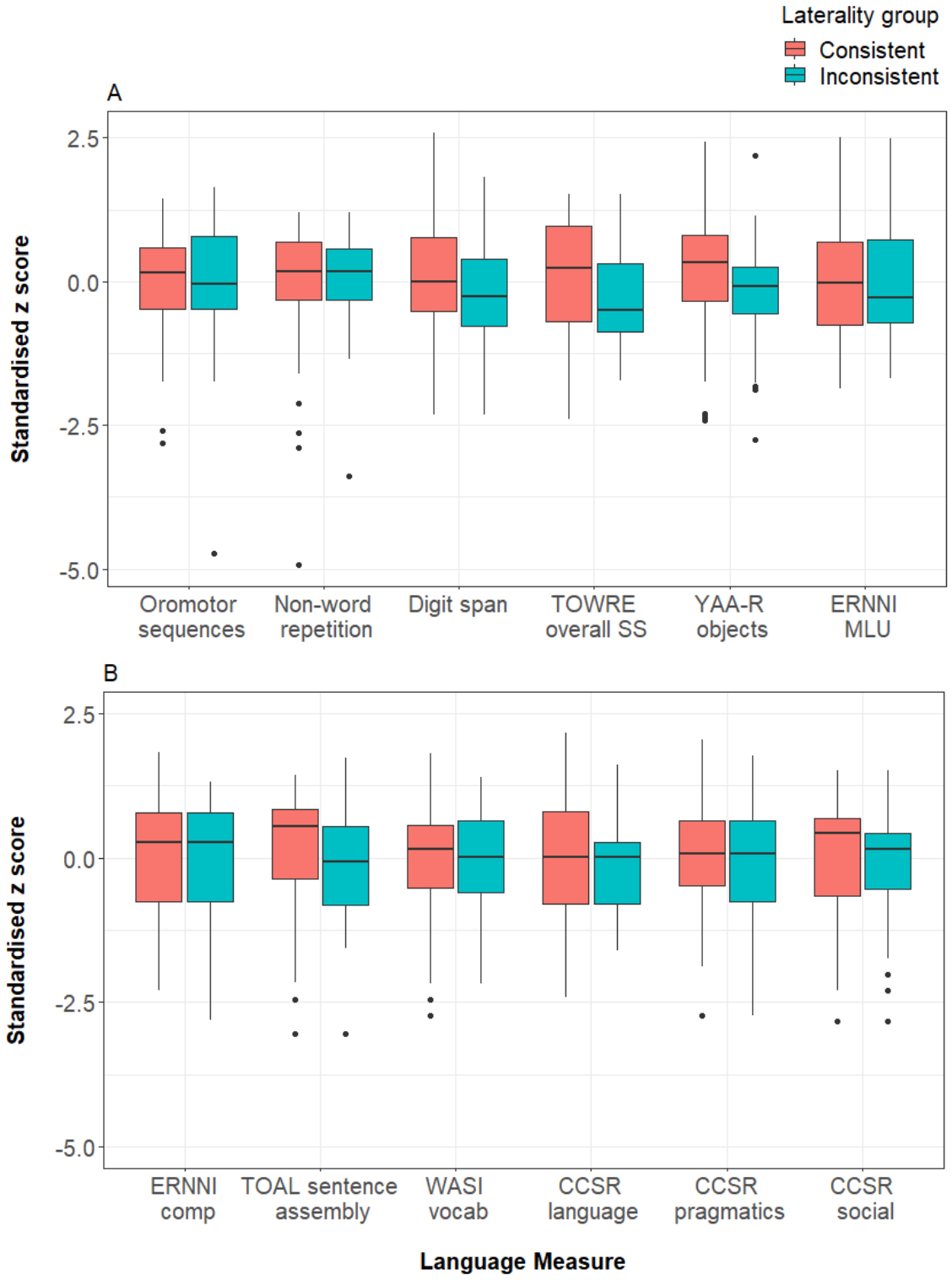


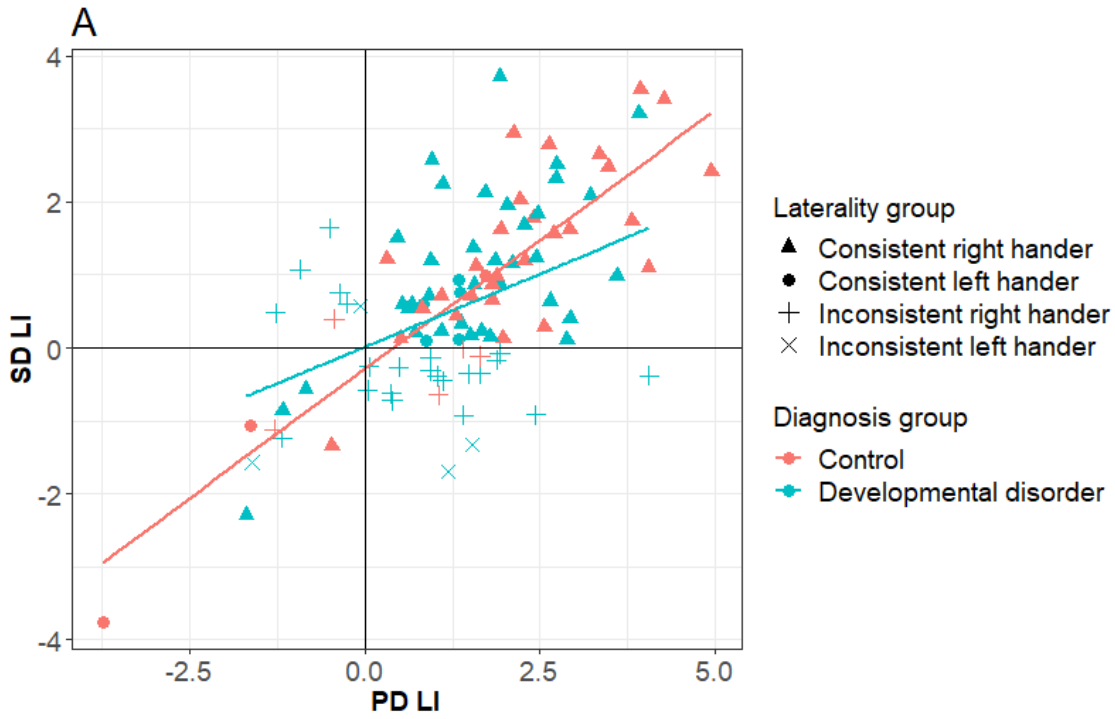
Figure 5.6: Laterality group language performance. *Standardised scores on each of the 12 language measures in the consistent and inconsistent laterality groups. The first six measures are shown in panel A, and the last six measures in panel B. Dark lines indicate medians, dots indicate outliers.*

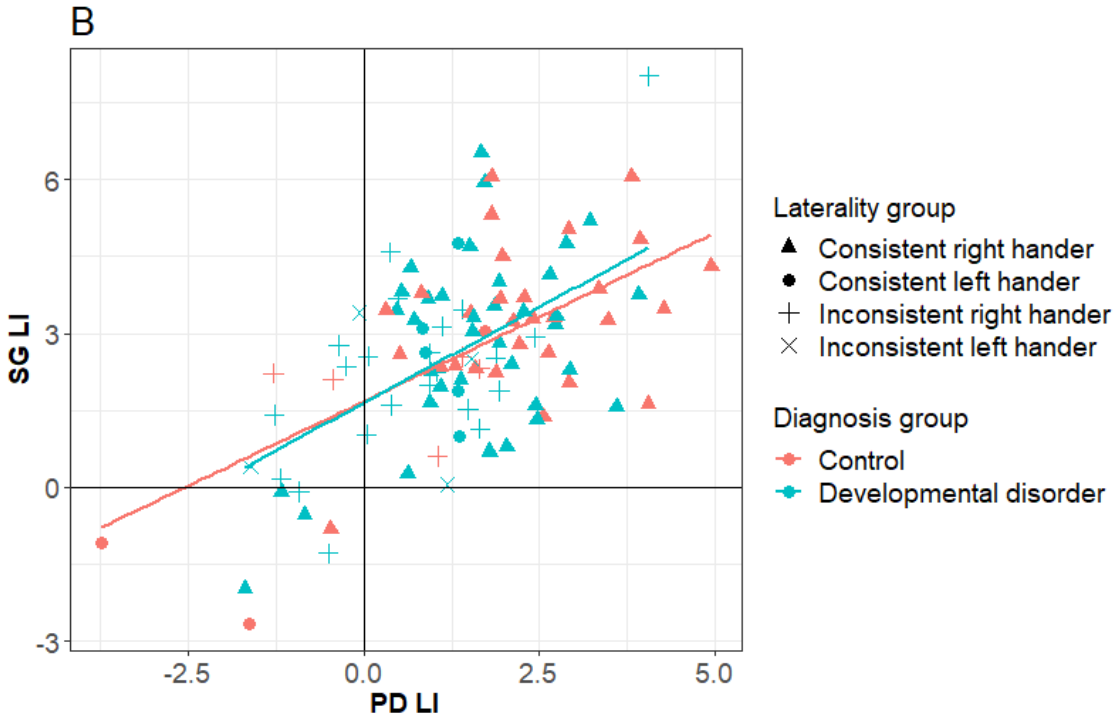
5.3.4 ***Non-preregistered exploratory analyses***

5.3.4.1 Correlations between task LIs

To further explore patterns of inconsistent lateralisation across the tasks, scatter plots were generated showing the relationship between LIs from the different tasks (see Fig. 5.7). Symbols classify participants in terms of consistency of laterality on fTCD, as well as handedness, and colour identifies whether each case is control or developmental disorder. Discrepant laterality between two tasks is indicated by the presence of points in the upper left and lower right quadrants of these plots. Most striking is the absence of points in the lower right quadrant of the PD-SG plot, indicating that no individuals demonstrated leftward lateralisation for PD but rightward lateralisation for SG.

To investigate relationships between laterality indices across the three tasks, Pearson's correlations were run for pairs of tasks for the developmental disorder and control groups separately. These can be found in Table 5.4. Using a Bonferroni corrected p value of 0.008 (0.05/6), all correlations were significant except that between sentence generation and semantic decision for the impaired group. Further, for the task pair phonological decision and semantic decision, inspection of the 95% confidence intervals for the Pearson's r statistics in the two groups revealed that these did not overlap. This indicates that the developmental disorder group had a weaker PD-SD correlation than the control group. In other words, the developmental disorder participants showed higher within-individual variability for this pair of tasks than controls. This pattern can be seen in the spread of data points around the regression lines for the two groups on these tasks in Figure 5.7A.


A
SD LI
PD LI
4
2
0
-2
-4
-2.5
0.0
2.5
5.0
Laterality group
Consistent right hander
Consistent left hander
Inconsistent right hander
Inconsistent left hander
Diagnosis group
Control
Developmental disorder


B
SG LI
PD LI
6
3
0
-3
-2.5
0.0
2.5
5.0
Laterality group
Consistent right hander
Consistent left hander
Inconsistent right hander
Inconsistent left hander
Diagnosis group
Control
Developmental disorder

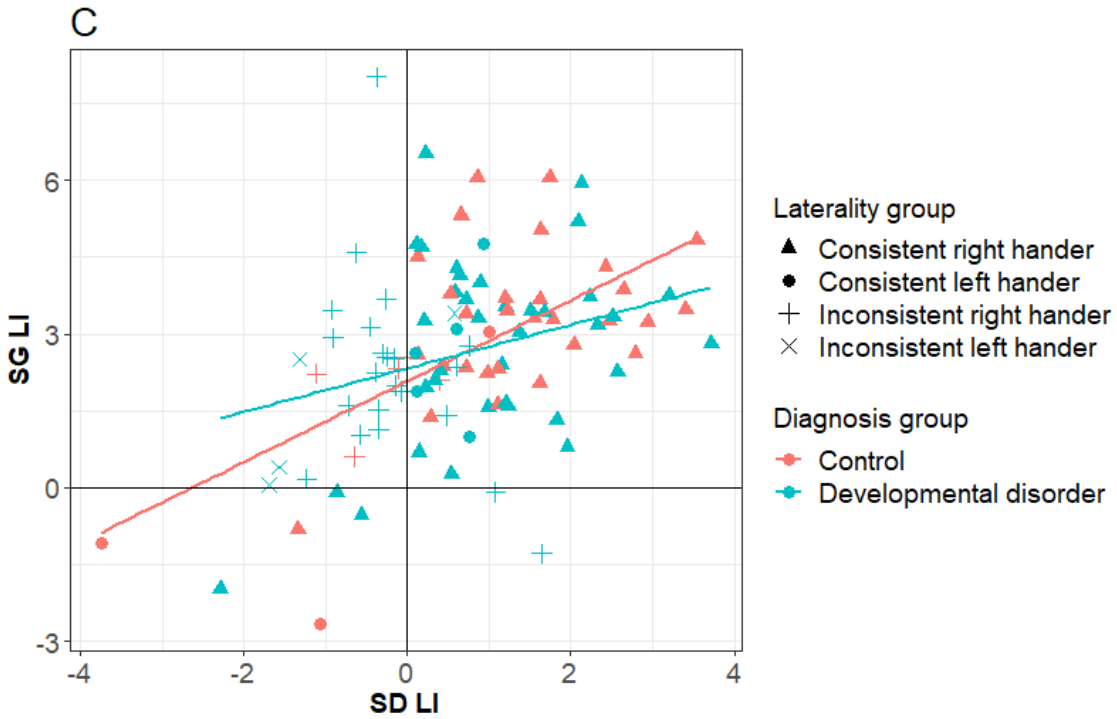


Figure 5.7: Scatter plots showing relationships between pairs of task LIs*: (A) Phonological decision and semantic decision, (B) Phonological decision and sentence generation, and (C) Sentence generation and semantic decision. Regression lines are fit separately for each group (control versus developmental disorder). Symbols distinguish participants in terms of handedness and consistency of laterality on fTCD; colour identifies whether each case is control or developmental disorder.*

Table 5.4: LI correlations in diagnosis groups. *Pearson's correlations between tasks for developmental disorder and control groups. PD = phonological decision, SD = semantic decision, SG = sentence generation.*

	Dev. disorder group r	Confidence interval	Control group r	Confidence interval
PD_SD	0.437	0.22-0.61	0.854	0.73-0.92
PD_SG	0.537	0.34-0.69	0.636	0.39-0.79
SG_SD	0.282	0.04-0.49	0.632	0.39-0.79

5.3.4.2 Consistency of lateralisation in the two groups

A summary of the numbers in each laterality group from the developmental disorder and control groups is given in Table 5.5. A chi squared test found that the frequencies from the two diagnosis groups in each laterality group were significantly different from those that would be expected by chance (χ^2 = 7.29, p = .007, df = 1). This demonstrates that the developmental disorders group contains a significantly higher frequency of individuals with inconsistent lateralisation than the control group.

Table 5.5: Frequencies in the two laterality groups from each diagnosis group.

Group	Consistent laterality	Inconsistent laterality
Developmental disorder	41	26
Control	32	5

5.3.4.3 Consistency in lateralisation over time

To investigate whether the group with developmental disorders demonstrated inconsistency in lateralisation from trial to trial within performance of a single task, the two groups were compared on a metric of lateralisation consistency across trials for each task individually (see *5.2.4.2 Non-preregistered exploratory analyses*). Using a Bonferroni corrected p value of .0167, no significant group differences were found for SD or SG; however, a significant group difference was found for PD (t(73.35) = 2.69, p = .009), with the developmental disorder group demonstrating reduced consistency in lateralisation across trials (see Fig. 5.8).

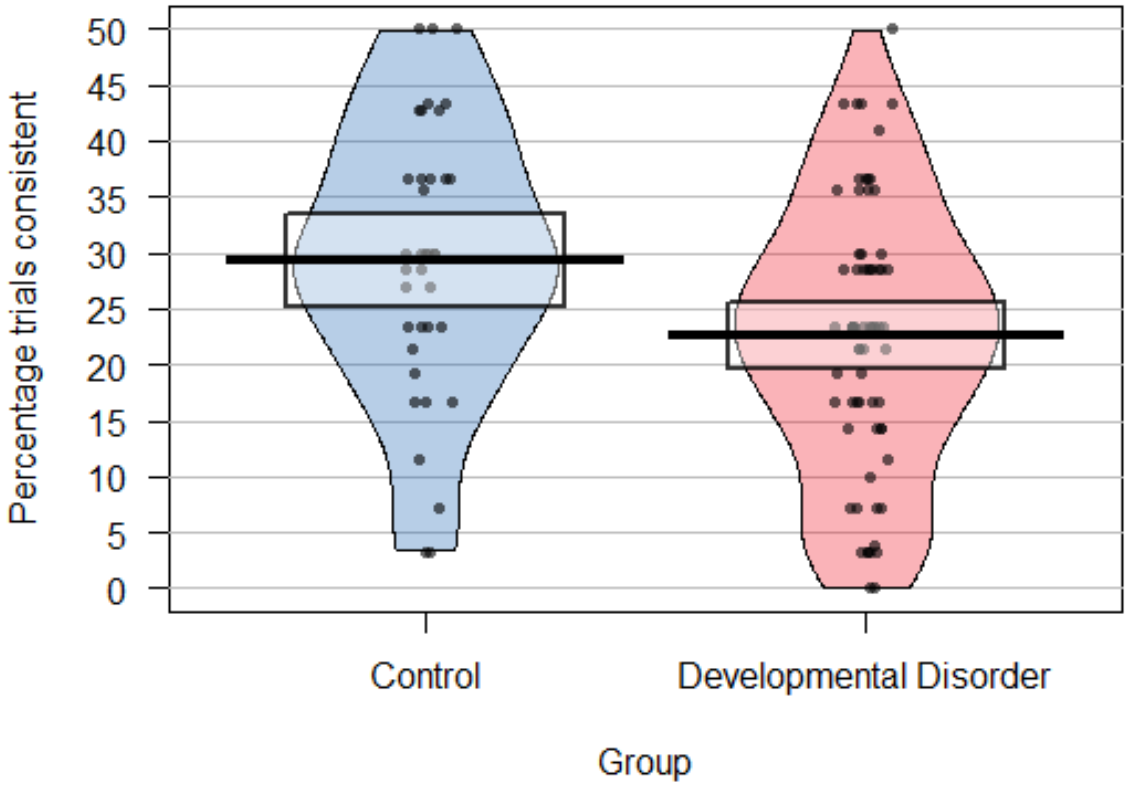


Figure 5.8: Consistency in lateralisation across trials. *Consistency in lateralisation over trials within the phonological decision task, for control and developmental disorder groups. Consistency in lateralisation was calculated as the absolute difference between 50 and the percentage of trials lateralised to the left. Higher values thus indicate greater consistency.*

5.3.4.4 Comparing strength of lateralisation

Previous literature on the relationship between language ability and lateralisation has focused on comparing strength of laterality between impaired and control groups on a single measure of lateralisation. Although the hypothesis that the relevant aspect of lateralisation affected in developmental disorders may be consistency rather than strength was not supported, the data was also analysed in a more conventional way to look at potential group differences in laterality strength. To do this, multi-level modelling was used to compare strength of laterality between developmental disorder and control groups across the three

laterality tasks. LI values (calculated from the mean method) in each group across the three tasks are shown in Figure 5.9.

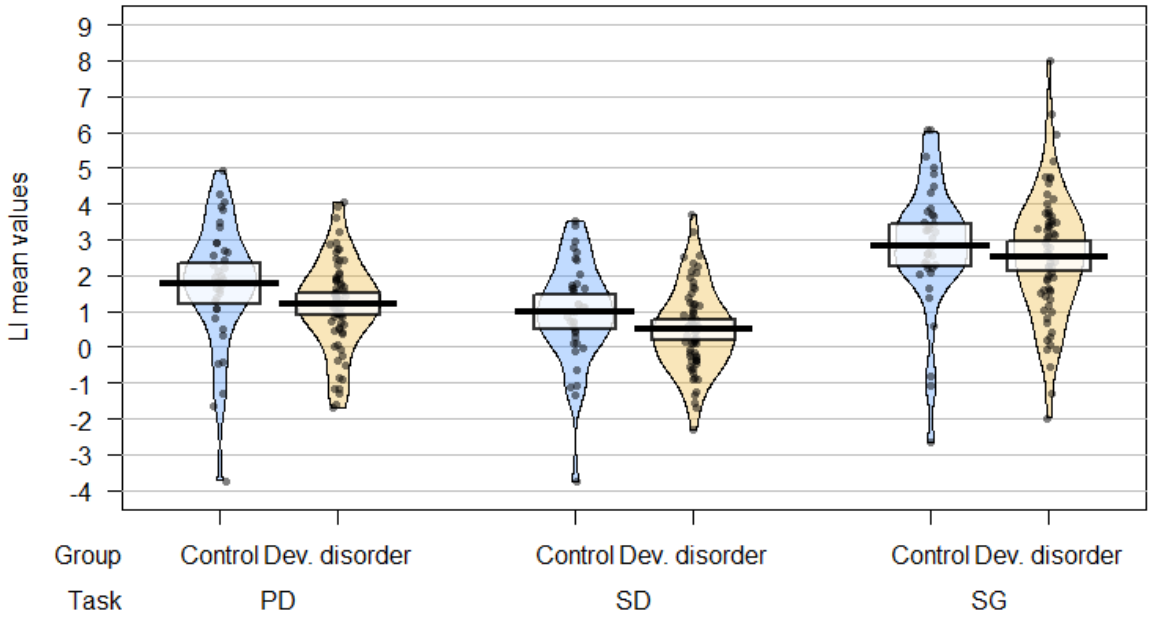


Figure 5.9: Strength of laterality. *LI values (calculated using the mean method) in developmental disorder and control groups for phonological decision (PD), semantic decision (SD) and sentence generation (SG) tasks.*

A multi-level mixed model analysis was conducted using the *lme4* package in R (Bates, Mächler, Bolker, & Walker, 2015). This entered group and task as fixed effects (without interaction) and participant as a random effect into a random intercept, fixed slope model. This model found no significant effect of group on LI values (t = 1.753, p = .083, df = 102). This did however find a significant effect of task; using PD as a reference condition, LIs were found to be significantly lower for SD (t = -5.08, p < .001, df = 206), but significantly higher for SG (t = 8.60, p < .001, df = 206).

5.4 Discussion

This study aimed to test the hypothesis that inconsistent lateralisation across different language tasks within individuals is associated with impaired language skills. This hypothesis was not supported; a multivariate test found no significant difference between consistent and inconsistent laterality groups across the 12 language measures. There was evidence, however, of increased within-individual variability in lateralisation in individuals with developmental disorders compared to controls. The developmental disorder group showed an increased incidence of inconsistent laterality patterns and weaker correlations between task laterality indices compared to the control group. Further, they demonstrated reduced consistency in lateralisation from trial to trial within the phonological decision task. They did not however show significantly weaker laterality than the control group in a multi-level modelling analysis with the three tasks. Overall, the results do not support the idea that poorer language functioning is associated with either increased inconsistency or reduced strength of laterality across language tasks. However, they do suggest that having a developmental disorder may be associated with increased likelihood of inconsistent lateralisation for different language processes.

5.4.1 *Lateralisation and language impairment: Direction of causality*

It has been suggested that atypical lateralisation for language may constitute an endophenotype for language impairment which mediates the relationship between genetic risk and expression of impaired language abilities. Although increased incidence of atypical laterality patterns has been associated with developmental disorders of language, generally atypical laterality itself does not appear to be associated with poorer language functioning (Bishop et al., 2014;

Knecht et al., 2001). The current data suggest that even when considering atypical laterality on a multidimensional scale (i.e. considering laterality across multiple language domains), atypical patterns manifest as inconsistent lateralisation across language tasks are not associated with impaired language. This suggests that even in cases where the language network is more distributed across the two hemispheres, typical language functioning can be maintained. That is, the conjectured increased demands on interhemispheric integration due to a more diffuse language network does not (at least not by itself) constitute a barrier to development of good language skills. This contradicts theories of the evolution of a lateralised brain which argue for the importance of sub-processes within a domain being located in the same hemisphere for optimal functioning (Kosslyn, 1987; Ringo et al., 1994).

The data did suggest however that the presence of a developmental disorder may increase risk for a more diffuse language network across the hemispheres. The developmental disorder group demonstrated a significantly increased incidence of inconsistent laterality patterns than would be expected by chance. Furthermore, correlations between task laterality indices tended to be weaker than in controls; in particular, correlations between LIs for phonological decision and semantic decision were significantly lower in the developmental disorder group than the control group. This suggests that this group showed greater within-individual variability in lateralisation for different language sub-processes. This provides support for the idea that in fact it may be the experience of having a developmental disorder that increases the likelihood of development of an unusual organisation of the language network, possibly due to attempts at compensatory reorganisation during development. This neuroplasticity model

was discussed by Bishop (2013); in this model, lateralization has no causal implications for language, but language impairment itself has consequences for how the brain develops. This model can reconcile the paradoxical findings of increased incidence of 'atypical' laterality patterns in developmental disorders on the one hand, and no significant differences in language skills between individuals with 'typical' and 'atypical' laterality patterns on the other. This evidence suggests the need to reconsider models of the relationship between language impairment and language lateralisation, and the assumption implicit in much laterality research that typical lateralisation patterns are beneficial for language functioning.

A developmental model of lateralisation was put forward by Minagawa-Kawai, Cristià and Dupoux (2011) who suggested that lateralisation is driven initially by existing biases in lower level sensory processing preferences of the two hemispheres, but increases in strength with increasing language experience over the first few years of development. For example, their research has reported that processing of native phonemic contrasts becomes left lateralised over the first year of life whereas that of non-native contrasts remains bilateral (Minagawa-Kawai, Mori, Naoi, & Kojima, 2007); they attributed this to increasing experience and skill with native phonemic contrasts over this developmental time frame. In this model therefore, language experience has an important role to play in driving the establishment of a specialised left lateralised language network. One could extend this theory to predict that in cases of impaired language, reduced proficiency and slower rate of development of language skill may result in incomplete or weaker establishment of a specialised left hemisphere language system. Experience of impaired development of language skills could therefore

impact lateralisation either by triggering attempts at compensatory reorganisation, or in fact by failing to engage normal mechanisms that drive the establishment of left lateralised processing circuits for language.

5.4.2 *Patterns of laterality across language tasks*

The current study is distinct from previous research in that it considered the relevance of within individual variability in lateralisation across different processes to language functioning. This allowed for observation of different patterns of dissociations in laterality across the language tasks. By far the most common pattern of inconsistent laterality was left lateralisation for phonological decision (PD) and sentence generation (SG), but right lateralisation for semantic decision (SD). PD and SG showed particularly high concordance, with no cases of left laterality for PD but right laterality for SG. This may reflect the common engagement of subvocal rehearsal by both tasks; SG required subvocal rehearsal of a generated sentence while PD required covert naming of the two pictures in order to make the rhyme judgement. These tasks are therefore likely to engage left posterior inferior frontal cortex (pars opercularis), which is important for subvocal rehearsal (Price, 2012).

Conversely, the SD task is considered a temporal-lobe task, typically engaging bilateral regions in the temporal lobe in fMRI studies (see Bradshaw, Thompson, Wilson, Bishop, & Woodhead, 2017 for a review). A meta-analysis by Rice, Lambon-Ralph, and Hoffman (2015) suggested that while semantic knowledge is represented bilaterally in the anterior temporal lobes, access to this information during linguistic processing was more likely to be left lateralised. While performance of the SD task encouraged linguistic access to semantic concepts, performance did not require explicit naming of the pictures in the same way as

the PD task; it is possible therefore that variability in the strength of left lateralisation of this task may reflect individual variability in engagement of a linguistic-based strategy for performance of the task. The frequent dissociation of PD laterality from SD laterality is however particularly striking given that this pair of tasks was designed to be as closely matched on non-linguistic parameters as possible; both involved presentation of pairs of line drawings with identical presentation timings and response modes (yes/no key press). This suggests that laterality measurement with these tasks is not disproportionately affected by their non-linguistic parameters but is tapping the underlying language processes.

Individuals with developmental disorders were even more likely to dissociate on SD and PD tasks than controls; correlations between these tasks in the two groups were significantly different from each other, indicating a higher level of within-individual variability across these tasks in the developmental disorder group. This represents tentative evidence for greater variability in frontal/speech production versus temporal/lexico-semantic lateralisation in individuals with disorders affecting language. Alternatively, it is possible that this pattern corresponds to differences in the strategy used; individuals with developmental disorders may have been less likely to engage a linguistic strategy for performance of the SD task (i.e. one which involved covert naming of the pictures before extracting category information), instead performing the decision based only on the pictures. The meta-analysis from Rice et al (2015) would predict that performance of the task without access to linguistic information associated with the pictures would be more bilateral than if category membership was extracted using linguistic labels. In this way, it is possible that differences in task strategy may underlie differences between the groups.

5.4.3 *Inconsistency in lateralisation within a task*

Alternatively, the higher likelihood of dissociation of laterality across PD and SD tasks within the developmental disorders group may reflect the increased level of inconsistency in lateralisation across trials observed in this group during performance of the PD task. The developmental disorder group appeared to engage in more frequent switching between dominance of the two hemispheres during the PD task. This could result in a less stable LI when calculated across all trials within the task, increasing the likelihood of discrepant laterality when compared with that of other tasks. Although the developmental disorder group did not demonstrate significantly worse performance on the PD task relative to the control group, it is likely that this phonological task was particularly challenging for the high number of dyslexic participants in this sample. It is possible that in order to support a matched level of performance, these participants may have attempted to recruit additional right hemisphere resources in order to compensate for faulty left hemisphere phonological processing. This more frequent switching between hemispheres during performance of the PD task may therefore reflect a compensatory mechanism in the current highly functioning sample.

5.4.4 *Further investigations using fMRI*

A pertinent question is whether the current results from this fTCD study warrant further investigation in fMRI studies. Despite the lack of support for the main hypothesis of a relationship between inconsistent laterality and poorer language skills, several hypotheses do follow from patterns observed in the developmental disorders group that would require fMRI investigation. Of particular interest is the hypothesis of increased frontal/temporal dissociations in laterality in individuals

with developmental disorders; testing this requires the spatial resolution of fMRI. Further, the exploratory analysis of inconsistency in lateralisation over time suggests that during performance of a phonological task, individuals with developmental disorders may switch between reliance on left hemisphere phonological representations and compensatory recruitment of right hemisphere additional processing resources. It would be of interest to use fMRI to further investigate this hypothesis and specify which right hemisphere areas may be recruited be these individuals; in theory these could either be the right homologues of the left language areas, or domain-general areas.

The current study constitutes the first piece of research investigating the relationship between language impairment and lateralisation with reference to the multidimensionality of lateralisation across different language functions. As such, the use of the cheaper and more quickly administered fTCD method to a large sample was considered a useful first step in testing for interesting patterns that could be followed up by later more expensive and time intensive fMRI investigations. This work has enabled development of more targeted research hypotheses, but also suggested that the search for an association between inconsistent laterality patterns and impaired language may not be fruitful. Conversely, further examination of how patterns of laterality across the language network can vary in individuals with language-relevant developmental disorders could shed light on how changing language experience may interact with compensatory changes in brain organisation.

Overall, the results of the current work highlight the importance and utility of using batteries of language tasks for laterality measurement with either fTCD or fMRI; the current data demonstrates that measurement with only one task does not

provide a sufficient description of individual variability in laterality patterns, and so is not appropriate for investigation of the significance of such individual variability to language functioning.

5.4.5 *Limitations of the study*

This study tested a sample of high functioning adults with a range of developmental disorders. Although these disorders are typically associated with impairments in language skills, most participants were students enrolled on a university course, and thus did not demonstrate severe impairments. It could be argued that a more clinically impaired or a child sample would have provided a better context in which to test this hypothesis. However, it should be noted that a previous chapter examining the profile of language abilities in the current sample of adults did establish that this sample was significantly impaired across a range of language measures compared to control participants (*Chapter 4: 4.3 Aim 1*). Thus, the current failure to find evidence for a difference in language abilities between consistent and inconsistent laterality groups cannot be attributed to a lack of a range of language ability levels in the sample. However, when interpreting the laterality differences in the developmental disorder group, the unusual nature of this compensated sample should be taken into consideration.

As already discussed, the benefits of the fTCD method must be weighed against a number of limitations when compared to other imaging methods. Measurement of lateralisation with fTCD is not only limited by a lack of regional specificity, but is also constrained by the territory of the middle cerebral artery. Although this covers many key areas within the language network, one region outside this territory is the ventral occipitotemporal cortex (vOT), often termed the ‘visual word form area’. Interestingly, fMRI work by Tailby, Weintrob, Saling, Fitzgerald, and

Jackson (2014) in patients with epilepsy suggests this region may demonstrate important differences in laterality according to language abilities. They reported that laterality in vOT was important for distinguishing epilepsy patients with reading disorders from those without, with the former demonstrating greater bilaterality in this area. VOT has been demonstrated to co-lateralise with frontal expressive language regions in healthy controls (Seghier & Price, 2013; van der Haegen et al., 2012). It is possible therefore that dissociation of laterality of this posterior area from inferior frontal language areas may constitute a risk factor for reading impairment. Unfortunately, such dissociations would go undetected by the fTCD method since the MCA territory does not cover ventral occipitotemporal regions (Badcock & Groen, 2017; Markus, Pereira, & Cloud, 2016). FMRI would enable testing of the hypothesis that inconsistent lateralisation across inferior frontal and ventral occipitotemporal regions is associated with reading difficulties.

5.4.6 *Summary and conclusions*

This study did not find evidence that inconsistent laterality across different language tasks (indicative of a more distributed language network) was associated with poorer language skills as hypothesised. However, the data does suggest that such inconsistent profiles of language laterality may be more likely in individuals with developmental disorders. The results thus favour a neuroplasticity model of the relationship between language impairment and language lateralisation, in which it may in fact be experience of development in the context of a developmental disorder that results in changes in brain organisation for language. Overall, the findings highlight the importance of multi-task paradigms for laterality measurement, and suggest new hypotheses that should be followed up with the increased spatial resolution of the fMRI technique.

Chapter 6 Testing the interhemispheric deficit theory of dyslexia using the visual half-field technique

Abstract

A deficit in interhemispheric transfer has been proposed as a neuropsychological theory of dyslexia. The role of interactions between the hemispheres during word recognition is demonstrated in two behavioural effects robustly reported using the visual half-field paradigm. The well-established recognition advantage for right visual field (RVF) words over left visual field (LVF) words reflects the additional processing costs associated with callosal transfer of LVF word representations to the language specialised left hemisphere. Additionally, a further gain in recognition for bilateral presentation of a word is attributed to cooperative interactions between the hemispheres. These recognition advantages can therefore be seen as behavioural indices of the efficiency of callosal transfer. This study aimed to replicate the finding of an absence of the bilateral advantage in developmental dyslexia, previously reported by Henderson, Barca and Ellis (2007). 47 dyslexic and 43 control adult participants were tested, and no significant difference found in the size of their bilateral advantages. The data did however replicate the previous finding of an increased RVF advantage in right handed dyslexic participants caused by poorer accuracy for LVF words. This evidence provides partial support for the interhemispheric deficit theory of dyslexia, suggesting an impairment in the transfer of visual word information from the right to the left hemisphere during reading.

6.1 Introduction

Language and reading require the integration of processing across multiple distributed regions, both within and across the hemispheres. This requires efficient sharing of sensory inputs and processing outputs between the hemispheres, predominantly mediated by the corpus callosum, the main white matter pathway connecting the two cerebral hemispheres. This structure has been proposed to be key in the establishment of lateralised circuits and in mediating interhemispheric transfer (van der Knaap & van der Ham, 2011).

6.1.1 *Models of reading*

In the case of reading, as we scan a text or read a centrally fixated word, orthographic information that falls in the left visual field (LVF) will be initially projected to the right hemisphere, with right visual field (RVF) input sent to the left hemisphere (due to decussation of optic fibres at the optic chiasm); at some point, processing of the two visual half-fields must then be integrated for whole word recognition, requiring callosal crossings. Further, while lower level visual processing is carried out bilaterally, there is evidence that processing of the word form becomes left lateralised at around the level of the left ventral occipitotemporal cortex (vOT); this area of cortex in the left hemisphere has been shown to respond with high selectivity to visual word forms presented at any location in the visual field (Cohen et al., 2002).

In their model of word recognition, Ellis et al., (2009) proposed that efficient recognition of a word is best supported when brain regions at the letter level (left middle/inferior occipital cortex) and word level (left mid fusiform gyrus i.e. vOT) engage in a state of mutual interaction. Ellis et al., argued that this kind of interaction could only occur between regions within the same hemisphere, and

not between hemispheres; hence in order for LVF word input to engage such interactive processing it must first be transferred across the corpus callosum to left vOT. Similarly, Price and Devlin (2011) argue that reading a word requires interactions between occipital, vOT and phonological language areas within the left hemisphere; the latter send top-down predictions to vOT based on learned mappings between orthography and phonology. A schematic of these models is given in Figure 6.1. Overall, visual word form input in the LVF initially projected to the right hemisphere must therefore be transferred to the left hemisphere for such higher order processing to allow efficient reading of the word. In this way, skilled reading can be seen to depend heavily on interhemispheric interactions and callosal transfer.

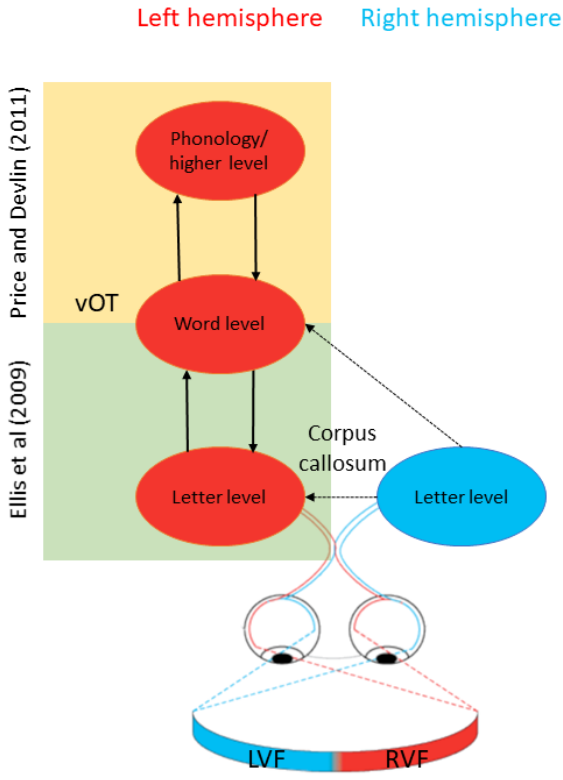


Figure 6.1: Models of word recognition*. Schematic of the Ellis et al., (2009) and Price and Devlin (2011) models that propose interactions between different levels in the word recognition process within the left hemisphere. Ellis et al. argue that word recognition is*

supported by interactions between letter and word level representations. Price and Devlin propose that interaction is required between occipital (visual), occipitotemporal (orthographic) and higher-order (e.g. phonological) regions.

6.1.2 *The visual half-field paradigm*

The role of the two hemispheres in word recognition has been investigated behaviourally using the visual half-field (VHF) paradigm. This task requires participants to report words that are briefly presented to the right or left visual hemifield. This has been used to demonstrate that words presented to the RVF are reported faster and more accurately than words presented to the LVF (Hellige, 1993). This RVF advantage has been attributed to the fact that the representation of a word presented to the LVF/right hemisphere has to be relayed to the left hemisphere via the corpus callosum in order to be processed by left lateralised language areas, decreasing efficiency of processing (Zaidel, Clarke, & Suyenobu, 1990). This interpretation is supported by research studying individual differences in language lateralisation. Although the majority of individuals show the typical pattern of left lateralisation for language, a minority (around 5%) show the reverse pattern of greater engagement of the right hemisphere during language processing (Knecht et al., 2000). Critically, there is evidence that these individuals with atypical language lateralisation as measured by fMRI also demonstrate an atypical LVF advantage in VHF tasks, suggesting that this measure is indeed sensitive to an individual's underlying organisation for language (Hunter & Brysbaert, 2008; van der Haegen et al., 2011).

A further effect that has been reported in visual half-field tasks is the phenomenon in which simultaneous presentation of identical word stimuli to the two visual hemifields results in significantly improved processing (increased accuracy and

faster reaction times) compared to when only the RVF is stimulated (Hellige, 1993). Different accounts of the cause of this so called ‘redundant bilateral advantage’ differ in whether they assume the need for interhemispheric integration. One type of explanation draws on ‘horse race’ models (Miller, 1982), and suggests that bilateral stimulation engages independent processing systems in each hemisphere, with recognition corresponding to whichever processor reaches the solution first. In this way, enhanced performance for bilateral presentation is due to probability summation whereby on average the operation of two independent processors will reach recognition faster than the operation of a single lateralised processor. This account has various shortcomings however. It cannot readily account for the robust finding that the bilateral advantage is sensitive to lexicality; unlike the RVF advantage, this effect is absent for pseudo-words (Mohr, Endrass, Hauk, & Pulvermüller, 2007; Mohr, Pulvermüller, Cohen, & Rockstroh, 2000). It also assumes that the two hemispheres have equal word recognition capabilities, a view which ignores the influence of left lateralised language circuits in aiding word recognition. In theory, race models can allow for a bias towards faster processing in one hemisphere than the other, as long as the probability distributions of time to word recognition for the two hemispheres overlap; this allows for the possibility that on some occasions, operation of both hemispheres on the same visual word input will result in faster recognition (by the right hemisphere) than if the left hemisphere were operating alone (see Fig. 6.2).

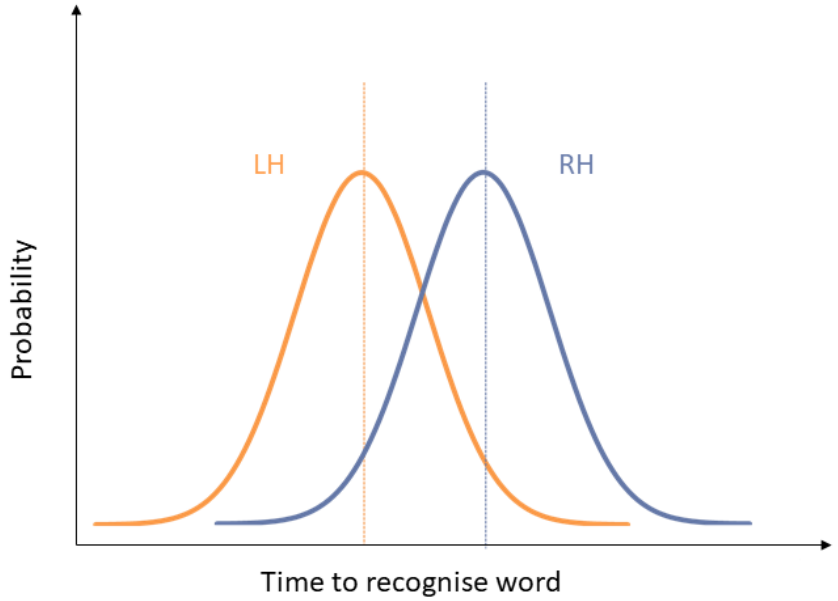


Figure 6.2: Pictorial illustration of race models of the bilateral advantage. *The overlap between the timing distributions of the two hemispheres means that on average, operation of both simultaneously increases the probability of a faster recognition time than operation of one hemisphere alone.*

A different model has been proposed by Mohr and colleagues, which builds on Hebb's assembly theory (Hebb, 1949). This model proposes that familiar stimuli are represented by 'transcortical cell assemblies' that consist of neurons across both hemispheres. Bilateral presentation will speed ignition of the entire assembly and so result in more efficient word recognition. This account posits a slightly greater role for interhemispheric integration, as this will be needed to allow spreading activation across the whole bilateral cell assembly. The assembly itself is assumed to have graded laterality, but Mohr et al. argue that both hemispheres possess capabilities for lexical processing (Mohr et al., 2007).

A further distinct view is put forward by Henderson et al., (2007). Their discussion of the causes of the bilateral advantage is rooted in the assumption that left visual field processing in the right hemisphere must be funnelled to the left vOT in order

for word recognition to occur. They argue that the benefit gained from bilateral stimulation relies on the left visual field input reaching left vOT while the right visual field input is still being processed; this results in a temporal summation of the inputs to improve recognition. Extending this to the Ellis et al. (2009) model of reading, this temporal summation may result in more successful inducement of a state of interaction between letter level and word level areas in the left hemisphere. Thus in this model, the bilateral advantage is heavily dependent on callosal transfer of left visual field input being efficient enough for this to reach left vOT at a point at which right visual field input is still being processed.

6.1.3 *Developmental dyslexia and interhemispheric processing*

Dyslexia is a developmental reading disorder that involves severe and persistent difficulties with reading and spelling. It is widely considered to involve a core deficit in phonological processing, in terms of mapping between orthographic inputs and phonological codes i.e. representations of speech sounds (Snowling, 2001). One neuropsychological theory of dyslexia proposes that these reading difficulties emerge as the result of disruption to the transfer of information across the corpus callosum, known as the interhemispheric deficit theory of dyslexia (Badzakova-Trajkov, Hamm, & Waldie, 2005; Monaghan & Shillcock, 2008; Orton, 1937).

It has been reported that individuals with dyslexia show deficits on tasks which require interhemispheric transfer, such as bimanual coordination and tactile transfer tasks (Daini et al., 2018; Moore, Brown, Markee, Theberge, & Zvi, 1995). Structural alterations of the corpus callosum have also been reported by multiple studies, although this evidence is not always consistent and likely depends on callosal sub-region (for a review, see Paul, 2011). An interesting parallel has been

described between dyslexia and the types of impairments shown by patients with agenesis of the corpus callosum, who are reported to show similar difficulties with phonological processing such as in rhyming (Banich & Brown, 2000). This evidence all points towards the relevance of abnormalities in interhemispheric transfer to the phonological and reading deficits shown in dyslexia.

Monaghan and Shillcock (2008) put forward a connectionist model of reading which allowed simulation of the effects of callosal disconnection. By including separate sets of units to represent each hemisphere, disconnection of these two 'hemisphere' units was found to result in impaired performance of the model for word reading, particularly for exception words with irregular spellings. The authors interpreted this as a model of 'surface dyslexia', a proposed sub-type of dyslexia characterised by particular difficulty in reading exception words (Bailey, Manis, Pedersen, & Seidenberg, 2004; Castles & Coltheart, 1993). Whilst this model was successful at simulating this reading impairment pattern, the lack of asymmetry with respect to processing within the hidden units representing the left and right hemispheres means that this model does not predict visual hemifield differences, and so fails to demonstrate the RVF advantage robustly reported in human word recognition. Therefore, despite its success in linking interhemispheric connectivity with reading skill, further models need to be developed that can incorporate hemifield differences in word recognition.

6.1.4 *Predictions of models*

Predictions of different models discussed here are given in Table 6.1 for different variables of interest in a VHF paradigm. These models differ in their predictions for patterns of visual word recognition in typical reading and reading in the context of hemispheric disconnection.

Table 6.1: Model predictions. *Predictions of models of word reading for variables of interest in VHF paradigms.*

Model	Predicts RVF advantage?	Predicts bilateral advantage?	Predicts greater RVF advantage after hemispheric disconnection?	Predicts smaller bilateral advantage after hemispheric disconnection?
Ellis et al (2009)	Yes	?	Yes	?
Price and Devlin (2011)	Yes	?	Yes	?
Horse race models (Miller, 1982)	No	Yes	No/Depends on response output	No/Depends on response output
Summation/ hemispheric cooperation model (Henderson et al., 2007)	Yes	Yes	Yes	Yes
Monaghan and Shillcock (2009)	No	Yes	No	Yes

Of particular note from Table 6.1 is the distinction between race models and summation models in their predictions of the effect of hemispheric disconnection on the bilateral advantage. It has been argued that race models do not predict that sectioning of the corpus callosum would affect the presence of the bilateral advantage (Mohr et al., 2007); in effect, this model considers the dynamics of the word recognition process for bilateral stimulation to occur in the same fashion in the connected brain as in the split brain. In contrast, summation models explicitly predict that hemispheric disconnection should eliminate the bilateral advantage. Thus, the report of an absence of a bilateral advantage in a split brain patient with sectioning of the corpus callosum has been used to argue against race models (Mohr et al., 1994). However, it should be noted that race models can in fact

predict a reduction in the bilateral advantage with callosal disconnection if response output is spoken; then, even in cases where the right hemisphere reaches word recognition first, performance will be solely determined by the time taken for the left hemisphere to reach word recognition and generate spoken output.

6.1.5 *Aims of the current study: A replication*

It is of interest then that Henderson, Barca and Ellis (2007) reported a significantly reduced redundant bilateral advantage on the VHF paradigm in dyslexic adults. Compared to controls, dyslexic participants showed both poorer LVF accuracy (resulting in a larger RVF advantage) and a lack of the usual significant difference between accuracy for RVF words and bilaterally presented words. By contrast, when words were presented in top and bottom visual fields, dyslexics showed the same redundant advantage as controls, suggesting that their failure to benefit from duplicate presentation relies on the need for interhemispheric transfer between the left and right visual fields. These findings were interpreted as evidence for a dysfunction in callosal transfer in dyslexia, meaning that LVF input is less able to both support good recognition alone and to provide a facilitatory effect when presented simultaneously with RVF stimuli.

The aim of the current study is to replicate this finding of a reduced or absent redundant bilateral advantage in dyslexic adults relative to typical readers, in order to address questions as to the processes underlying typical and impaired reading. Further, it was planned to relate the sizes of the right visual field and bilateral advantages to measures of word and non-word reading ability.

6.2 Method

This study was pre-registered as a replication study on Open Science Framework prior to commencement of data collection; the registration form can be found at https://osf.io/6a35m/.

6.2.1 *Participants*

47 individuals with a diagnosis of dyslexia (mean age: 23.53 years, 7 left handed) and 43 individuals without any diagnosis of language and/or reading impairment (mean age: 21.87 years, 2 left handed) were recruited for this study. Participants were recruited predominantly from local institutes of higher education, either via disability services for recruitment of the dyslexic group or via more general advertisements for the control group. Inclusion criteria were English as a first language and an age of 18-45 years. Exclusion criteria were significant hearing loss, history of neurological disease, head injury or epilepsy. Individuals recruited for the dyslexic group all had a confirmed diagnosis of dyslexia typically from an educational psychologist or the disability services at their university, and most were receiving special considerations during their studies. Since there were unequal levels of left handers in the two groups, subsequent analyses were also run with right handed participants only.

6.2.2 *Materials and Method*

Participants were given the Test of Word Reading Efficiency (TOWRE, Rashotte, Torgesen, & Wagner, 1999) to obtain scores for word and non-word reading. This test requires participants to read as many words/non-words as they can within a period of 45 seconds. Participants also completed tests 1 and 2 of the Cattell Culture Fair test (Cattell & Cattell, 1973) to give a measure of non-verbal IQ. These were administered to all participants except one dyslexic participant whose

VHF accuracy measures were nevertheless included in the data analysis. Mean performance on these tasks and mean age for each group are given in Table 6.2. The dyslexic group demonstrated significantly poorer performance on the TOWRE than controls (Mann Whitney $U = 302.5$, $p < .001$) and also unexpectedly on the Cattell (Mann Whitney $U = 637.00$, $p = .005$). There was also a significant difference in age between the groups (Mann Whitney $U = 673.00$, $p = .009$). In subsequent analyses, it was explored whether non-verbal abilities or age may have influenced the results, but this possibility was ruled out.

Table 6.2: Characteristics of sample. *Mean age and performance of each group on measures of reading and non-verbal IQ. Values for the dyslexic group on these standardised tests are shown for n = 46 since one dyslexic participant who participated in the VHF task did not complete these.*

Measure	**Controls**		**Dyslexics**	
	Mean	SD	Mean	SD
Age (years)	21.87	3.89	23.53	4.93
TOWRE word reading standard score	105.19	9.45	90.28	13.87
TOWRE non-word reading standard score	109.23	9.87	92.00	13.84
TOWRE overall standard score	109.21	10.04	89.30	15.45
Cattell non-verbal IQ (score out of 24)	21.00	2.15	19.57	2.36

For the visual half-field task, the same method and task design as Henderson et al., (2007) was employed, with minor adjustments. Stimuli were 180 five-letter

words selected from the CELEX database; these consisted of lists of words combined from across multiple experiments reported in the Henderson et al. paper in order to increase the number of trials in the task from 120 to 180. This was motivated by recommendations made by Hunter and Brysbaert (2008) who argued that at least 150 trials are required in order to yield reliable results with the visual half-field paradigm. 18 practise word stimuli were also used. For the main part of the task, 60 words were allocated to each visual presentation condition; right visual field (RVF), left visual field (LVF) or bilateral visual field (BVF) presentation. Each word was presented in upper case in Fixedsys font in size 18 point. They were presented on a 36.5 x 23 cm screen with a resolution of 1024 x 768 pixels, at a viewing distance of 60 cm kept constant by the use of a chin rest. The visual angle from the outer edge of the word to the fixation point was around 4.3 degrees, and from the inner edge of the word to fixation around 1.4 degrees; hence, words were presented in parafoveal vision (defined as 1.5 to 5 degrees). The task was run using E-prime 1.1 software (Psychology Software Tools, Pittsburgh, PA) using a script identical to that used by Henderson et al., (2007).

A trial began with presentation of a central fixation cross for 400 ms. This was followed by presentation of the word stimulus/stimuli for 60 ms in one of the three presentation styles. For unilateral presentation trials, a series of 5 Xs were presented in the other visual hemifield in order to balance the visual presentation (see Fig. 6.3) Following stimulus presentation, participants were required to use the keyboard to type the previously presented word, and then press enter in order to start the next trial. Participants were instructed that they could correct any spelling or typing mistakes before pressing enter and to provide their best guess

if they were unsure of the word. Scoring allowed for spelling and typing mistakes such that obvious errors were still counted as accurate responses (e.g. THEIF instead of THIEF). Within an experimental session presentation conditions were pseudo-randomly interleaved; across participants the order of these was counterbalanced.

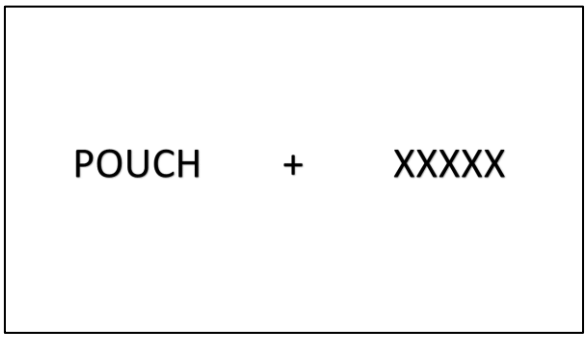


Figure 6.3: VHF stimulus display. *Schematic of stimulus display of target word and contralateral Xs. Shown for a LVF condition trial.*

6.2.3 ***Hypotheses and analysis***

6.2.3.1 Pre-registered analyses

It was predicted that, in line with the findings from Henderson et al. (2007), dyslexic adults would show a smaller difference between accuracy in the RVF and BVF conditions compared to typically developing adults; this finding would constitute a successful replication. To test this, for each individual the difference score between these two conditions was calculated, henceforth referred to as the RBA (redundant bilateral advantage). An independent groups t-test was then used to compare the magnitude of the RBA between the two groups. It was predicted that the size of this difference would be significantly smaller in the dyslexic group than in the typically developing group. This analysis deviates from the original analysis reported by Henderson et al., (2007). They used a 2x3

ANOVA to compare recognition accuracy between the two groups for RVF, LVF and BVF words, and reported a main effect of group, a main effect of condition, and a significant group by condition interaction, indicating that dyslexics showed a smaller bilateral advantage than controls. However, this interaction effect was associated with a very small effect size of r = 0.08, meaning that power calculations based on such an analysis with this effect size gave unrealistically high sample sizes of 200 plus participants. Therefore it was planned to use this simpler t-test analysis with a more focused hypothesis in order to increase power (see Lazic, 2017 for a discussion).

Based on an effect size of 0.957 for this simpler analysis calculated from the data reported on by Henderson et al. (2007), a power calculation using G*Power software demonstrated that 20 participants would be needed in each group to yield a power level of 0.9. However, since effect sizes can be overestimated in published studies particularly if they are under-powered (Button et al., 2013), the aim was to collect data on 30 participants in each group. This data was collected as part of a larger study, and by the end of the data collection period data was available with this task on 47 dyslexics and 43 controls.

6.2.3.2 Non-pre-registered analyses

Two further hypotheses were also tested that were not originally pre-registered. Firstly, it was predicted that dyslexics would show poorer LVF performance compared to controls; this constitutes an extension of the replication of the findings of Henderson et al., (2007) who reported that dyslexic LVF accuracy but not RVF accuracy was significantly lower than in controls resulting in a greater RVF advantage. Therefore this was tested using t-tests to compare LVF accuracy and RVF accuracy between groups. All t-tests were Bonferroni corrected; since

three t-tests were conducted a corrected *p* value of .016 was used. To test more directly for an increased RVF advantage in dyslexia this analysis was also followed up by a multi-level mixed modelling analysis to test for a significant interaction between group and condition. Two random intercept models were compared; a null model with fixed effects of group and condition (with no interaction) and a random effect of participant; and a full model with the same fixed and random effects but with an interaction between group and condition. A likelihood ratio test was used to obtain a p-value for the comparison between the full model and the null model.

Secondly, it was predicted that the size of the RBA and RVF advantage would correlate with measures of word and non-word reading performance as assessed by the TOWRE. On the assumption that increased efficiency of interhemispheric transfer as indexed by a small RVF and large RBA is beneficial for reading, a negative correlation was predicted for word and non-word reading scores with RVF advantage scores, but a positive correlation predicted with the size of the RBA.

While the Henderson et al. (2007) paper only included left handed participants, the current sample comprised both right and left handers since VHF data were collected as part of a larger study presented in another chapter of this thesis (*Chapter 5*). In order to rule out handedness as a potential confound in analyses, the group comparisons were therefore re-run excluding left handers.

6.3 Results

6.3.1 *Preregistered analysis: Replication of comparison of RBA in dyslexics and controls*

Means and standard deviations for each visual half-field condition can be found for each group in Table 6.3 along with their mean RBA and RVF advantage scores. Percentage accuracy for each condition is plotted in Figure 6.4 and given in Table 6.4 showing both the data reported on by Henderson et al. (2007) and the current study. Five controls and one dyslexic participant in the current sample failed to demonstrate the typical RVF advantage (LVF and RVF scores were either identical or greater for LVF than RVF). Handedness could not explain this individual variability in the presence of the RVF advantage, since all left handers demonstrated a positive RVF advantage.

RBA scores for each group can be seen in Figure 6.5. An independent samples t-test found no significant difference in the size of the RBA between groups, $t(88) = .859$, $p = .393$. Thus, this study did not replicate the finding of a reduced bilateral advantage in a dyslexic sample compared to controls. This result was unchanged if only right handers were included in the analysis. Since neither age nor Cattell score correlated with RBA scores, group differences in these variables are unlikely to have affected this result.

Table 6.3: VHF task performance. *Mean accuracy (and SDs) out of 60 on VHF measures for the Control and Dyslexic groups*

Group	LVF accuracy	RVF accuracy	BVF accuracy	RVF advantage	Redundant bilateral advantage
Controls	34.65 (SD = 10.49)	47.02 (SD = 7.41)	51.86 (SD = 5.51)	12.37 (SD = 11.39)	4.84 (SD = 5.21)
Dyslexics	27.74 (SD = 11.64)	43.91 (SD = 7.16)	47.74 (SD = 7.12)	16.17 (SD = 9.45)	3.83 (SD = 5.85)

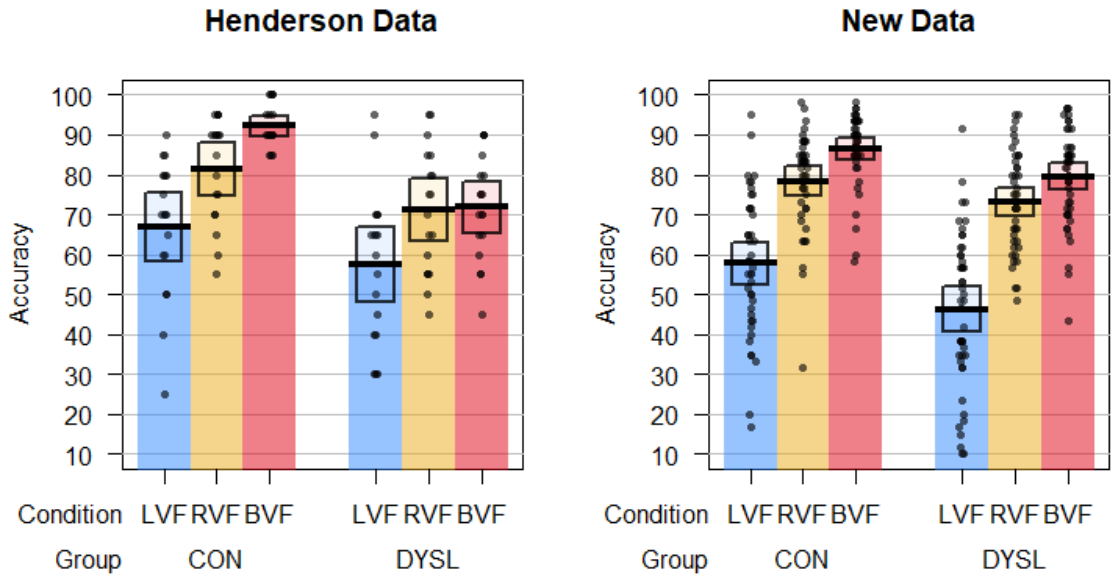


Figure 6.4: Performance on VHF conditions. *Accuracy data reported from the original study by Henderson et al. (2007) and the current study. Plots show individual and mean accuracy scores (as percentage accuracy) for each visual hemifield condition in the two groups. Boxes indicate 95% confidence intervals. Note that data for the Henderson study is given for n = 18 in each group, since data for the full groups reported on in the paper were not available. CON = control group; DYSL = dyslexia group.*

Table 6.4: Percentage accuracy on VHF task. *Percentage accuracy in the two groups across the three visual half-field conditions for Henderson et al. and the current study.*

	Dyslexics			**Controls**		
	LVF	RVF	BVF	LVF	RVF	BVF
Henderson et al (n = 44)	53%	72.7%	72%	68%	80.7%	92%
Current study (n = 90)	46%	73%	79.6%	57.7%	78.4%	86.4%

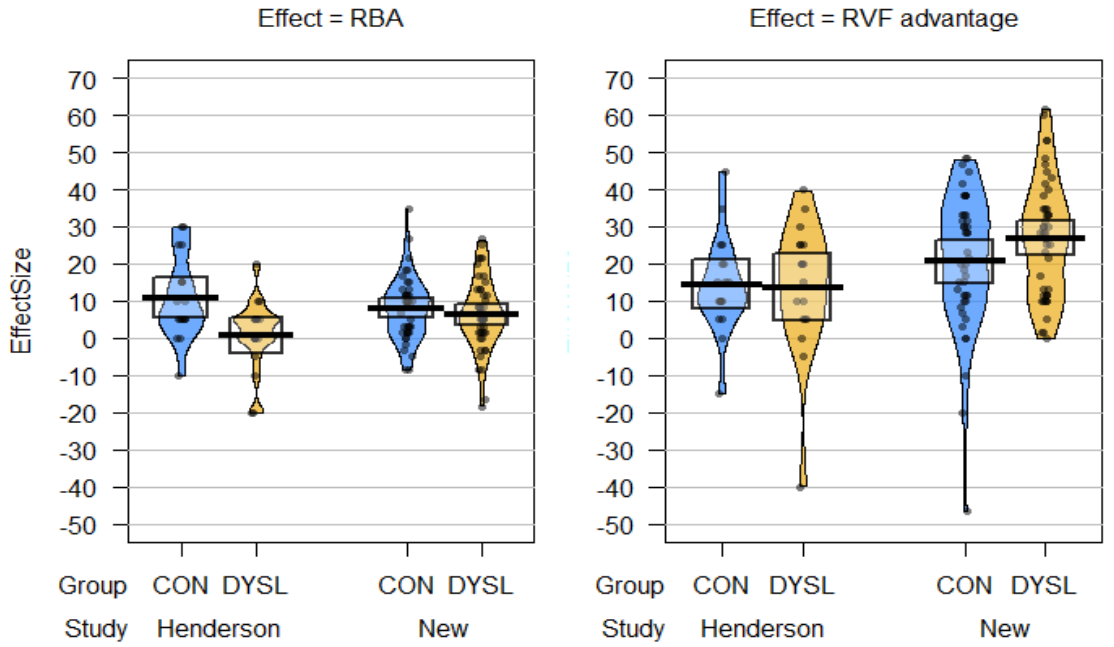


Figure 6.5: Bilateral and RVF advantages. *Pirate plot showing mean and individual data points for redundant bilateral advantage (RBA) and RVF advantage scores in the two groups, for the original Henderson et al. (2007) study and the current study (New). Shaded boxes indicate 95% confidence intervals. Note that data for the Henderson study is given for n = 18 in each group, since data for the full groups reported on in the paper were not available. CON = control group; DYSL = dyslexia group.*

6.3.2 *Non-pre-registered exploratory analyses*

A non-preregistered analysis was conducted to replicate a second analysis done by Henderson et al., (2007) comparing LVF and RVF accuracy between groups, with the hypothesis that the former and not the latter will be significantly different. An independent t-test found a significant difference between groups for the LVF condition, with dyslexics showing lower accuracy than controls ($t(88)$ = 2.95, p = .004, d = 0.62). Since RVF accuracy data was found to be non-normal, this was compared between groups by means of a Mann-Whitney U test. This did not find a significant difference between groups at the Bonferroni corrected p value (Mann-Whitney U = 720.5, p = .019, r = 0.247). Therefore this study replicated Henderson et al's (2007) finding of poorer performance of dyslexics for LVF presented words, which they interpreted as evidence of an increased RVF advantage in the dyslexic group compared to controls.

To test more directly for an increased RVF advantage in dyslexia, a multi-level mixed model analysis was used to test for a significant interaction between group (control versus dyslexic) and condition (RVF versus LVF) on accuracy scores. A likelihood ratio test was used to compare a null model with fixed effects of group and condition (with no interaction) and a random effect of participant, with a full model with the same fixed and random effects but with an interaction between group and condition. This found that the full model did not provide a better fit to the data, ($\chi 2(1) = 3.00$, p = 0.083). Re-running the same analysis with data from only right handed participants however did find a significant group by condition interaction ($\chi 2(1) = 4.71$, p = 0.030). For these participants, performance of the dyslexic group showed a greater difference between LVF and RVF conditions

than that of the control group, indicating a larger RVF advantage in the dyslexic group (see Fig. 6.6).

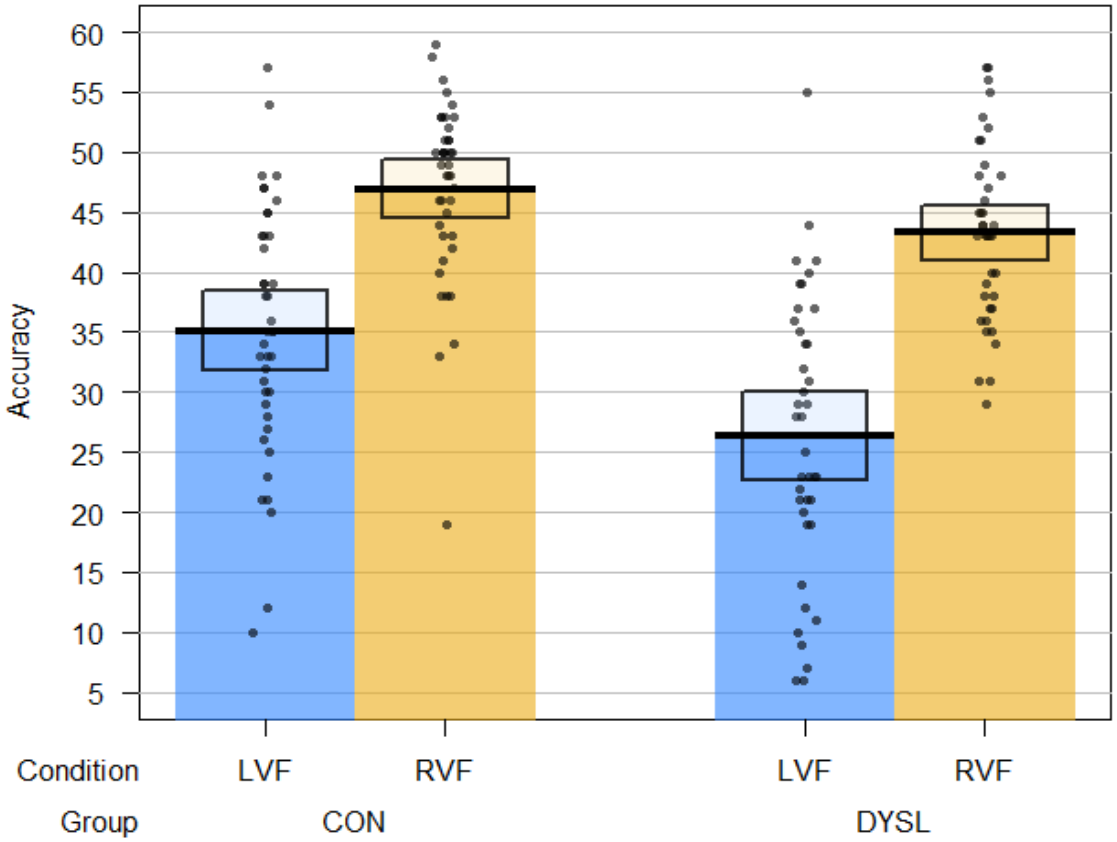


Figure 6.6: RVF advantage. *Performance in the RVF and LVF conditions in sample of right handed participants only. CON = control group, DYSL = dyslexia group.*

Further, exploratory analyses were run to investigate the relationship between TOWRE word and non-word reading standard scores and size of the RBA and RVF advantage. Pearson's correlations found no significant relationship between TOWRE word reading standard scores and either the RBA (r = .002, p = .983) or RVF advantage (r = -.102, p = .341). Similarly, no significant correlations were found between TOWRE non-word reading standard scores and either the RBA (r = -.03, p = .784) or RVF advantage (r = -.072, p = .503).

6.4 Discussion

This study aimed to use the visual half-field paradigm to replicate the finding of a reduced redundant bilateral advantage in dyslexic adults compared to controls. Analysis of the current sample failed to replicate this effect, with no significant difference between the size of the bilateral advantage in dyslexics and controls. The study did replicate the finding of an increased right visual field (RVF) advantage in dyslexics compared to controls, but only when excluding left handed participants. This was driven by significantly worse performance of the dyslexics on the left visual field (LVF) condition as previously reported in a right handed only sample by Henderson et al., (2007). Therefore, the current results provide partial support for the hypothesis of impaired interhemispheric transfer in dyslexia, but suggest that the nature of this impairment does not necessarily interfere with the bilateral advantage. Correlation analyses found that neither word nor non-word reading scores were significantly correlated with either RVF advantage or bilateral advantage scores.

6.4.1 *Poorer LVF accuracy in dyslexia*

The main pattern that was observed in the performance of the dyslexic group was a particular difficulty with words presented in the LVF. In contrast, performance for RVF words did not differ from that of controls. This increased RVF advantage in dyslexics appeared as a trend in the full sample and resulted in a significant group by condition interaction when left handers were excluded. This replicated Henderson et al's (2007) finding, who also reported an increased RVF advantage in dyslexia driven by poorer LVF performance in their sample of right handers. The inclusion of data from left handers in the current study may have dampened this effect in the full sample, due to a higher number of left handers in the dyslexic

group (7 versus two in the control group); this may have inflated the average dyslexic LVF performance. The replication of this pattern when restricting the sample to right handers to increase similarity to the sample reported on by Henderson et al. suggests that this is a robust effect in right handers.

This pattern of a specific difficulty with LVF words is consistent with the interhemispheric deficit theory of dyslexia. According to models of reading (Ellis et al., 2009; Price & Devlin, 2011), accurate recognition of LVF words requires transfer of this visual input across the corpus callosum to access left lateralised processing circuits involving reciprocal interactions between the letter, word and phonological levels. Impairment in the transfer of this LVF input across the corpus callosum would therefore prevent access to this more efficient mode of processing, meaning word recognition has to rely on more basic and perhaps more serial letter by letter processes in the right hemisphere (Ellis et al., 2009).

The idea that dyslexic readers struggle particularly with reading word forms present in their LVF is also supported by eye tracking research. Kelly, Jones, McDonald and Shillcock (2004) studied eye movements in dyslexics and controls during reading of sentences and found that dyslexics showed (i) more fixations to the left (i.e. at the beginning) of words and (ii) shorter fixation durations when fixating on the beginning of words than their centre. They interpreted this bias for more leftward fixations as a strategy employed by dyslexic readers to ensure most letter information in a word is directly projected to the left hemisphere in order to avoid the need for callosal transfer. Shorter fixation durations when fixating on such word beginning locations was further interpreted as evidence of reduced difficulty of processing, as opposed to cases where central fixations required integration of letter information about a single word across the

hemispheres. Such patterns in eye movements could constitute a compensatory mechanism for the disrupted transfer of LVF input to the left hemisphere as suggested by the current findings.

Interestingly, there is evidence that dyslexic participants may show the reverse pattern of heightened hemifield preferences when the task is right lateralised. Daini et al. (2018) measured performance of dyslexic children on an orientation matching task where a picture pair with either matching or discrepant orientations was presented to either the right or left visual field. In control participants, this task was found to involve a LVF advantage, indicative of right lateralisation. Their dyslexic participants demonstrated an increased LVF advantage, consistent with an impairment in the transfer of RVF input to the specialised right hemisphere. Interestingly, such a pattern has also been reported in split brain patients, who showed a larger LVF advantage for this orientation matching task (Corballis et al., 2010). Together with the present findings, this therefore suggests that dyslexia involves a general impairment in the transfer of visual information across the corpus callosum, regardless of the required direction or the stimulus domain.

6.4.2 *Failure to replicate a reduced bilateral advantage in dyslexia*

More puzzling however is why such a deficit in interhemispheric transfer suggested by poorer LVF performance was not apparent in performance for bilateral presentation in the dyslexic group. When interpreting their findings, Henderson et al., (2007) suggested that the absence of the bilateral advantage in dyslexics may be due to issues in the timing of interhemispheric transfer; that is, the LVF input may not reach the left hemisphere in time to summate with the RVF input and 'boost' processing. An alternative view would be that interhemispheric transfer is not slower but instead degraded such that the

representation of the LVF word form transmitted to the left hemisphere is of poor quality. Such a scenario was in fact induced in typical readers in an experiment in the Henderson et al., (2007) paper. Here, RVF input was presented simultaneously with a degraded version of the same word in the LVF in the form of a word fragment consisting of the first and last letters of the word (e.g. c--n for coin). Interestingly, this bilateral presentation was sufficient to induce a bilateral advantage in the controls; however, lateralised presentation of the word fragment alone did not allow for its successful recognition. In light of the current findings, one might suggest that the bilateral advantage demonstrated in the dyslexic group could be supported by a similar process of integration of degraded LVF input (due to faulty interhemispheric transfer rather than stimulus manipulation) with full word RVF input. That is, this LVF input can arrive in time to summate with the RVF input and result in a bilateral advantage, but is degraded in quality and so insufficient by itself to support accurate word recognition.

These two models of the nature of the deficit in interhemispheric transfer in dyslexia should be further tested against each other. The latency of interhemispheric transfer times (IHTTs) can be measured using ERPs to record the delay between a particular ERP component evoked in the contralateral hemisphere by a word stimulus and that evoked in the ipsilateral hemisphere. This method has been used to demonstrate a directional asymmetry in controls in which transfer of verbal material is faster from the right hemisphere to the left than vice versa (Brown et al., 1994). Using this method Martin, Thierry, Démonet, Roberts, and Nazir (2007) reported an absence of this asymmetry in dyslexic participants, who instead showed balanced latencies for transfer of visual word information in both directions. However, whilst overall IHTTs were reported as

‘marginally shorter’ in controls than dyslexics, this did not reach statistical significance in their small sample (10 participants per group) and therefore should not be interpreted as evidence for slower callosal transfer in dyslexia. Computational model simulations such as that developed by Monaghan and Shillcock (2008) could be further developed to test these two hypotheses. In their model, callosal transfer was completely abolished such that there was no communication between the two units of nodes representing the two hemispheres. It would be of interest to model callosal deficits in a more graded fashion; connections between the two ‘hemisphere’ neural networks could either be made slower (for example with longer connection lines or another layer of nodes), or pruned such that fewer connections remain resulting in incomplete transfer of information.

6.4.3 *Methodological differences between the studies*

The current failure to replicate the group difference in the bilateral advantage reported by Henderson et al., (2007) could simply indicate that this previous finding was a false positive which the current more highly powered study was able to avoid. However, it is important to examine potential differences between this study and theirs. Figure 6.4 and Table 6.4 reveal that performance of control participants across the three conditions was slightly reduced compared to that of the controls reported on by Henderson et al; further the current bilateral advantage was slightly smaller (around 8% versus 12%). This lower accuracy is puzzling given that the same task was administered using the same script to very similar groups of participants. The key difference between the procedures of the two studies was in the number of trials, with the current study increasing this from 120 to 180 in line with recommendations from Hunter and Brysbaert (2012). This

increase in the number of measurements taken would have further increased the power of the design to detect group differences.

The make-up of the current dyslexic and control groups were highly similar to those of the original study in age, student status and performance on the same measure of reading and non-word reading ability. The data did however show a higher level of individual variability in VHF accuracy scores than that reported by Henderson et al., (2007), with higher standard errors for each condition in both groups. Visual half-field data is often characterised by a high level of individual variability, which has led many to argue that behavioural methods are too noisy to provide reliable measures (Hunter & Brysbaert, 2008; Voyer, 1998). Despite this however, a recent study demonstrated that the majority of effects at the group level reported in the visual half-field literature do replicate in well powered studies (Brederoo, Nieuwenstein, Cornelissen, & Lorist, 2019). This suggests that visual half-field paradigms do yield robust and reliable group effects, despite their potential for high individual variability.

The current study did deviate from the methods of the original Henderson et al. study in its inclusion of left handed participants. However, removing these left handed participants did not affect the outcome of the group comparison of the bilateral advantage, and therefore this difference cannot explain the failed replication. The inclusion of left handed participants did however affect the comparison of the RVF advantage between groups; a significant condition by group interaction was only found when restricting the sample to right handers. This is likely attributable to the uneven levels of left handedness in the two groups; the increased number of left handers in the dyslexic group may have served to dampen this group difference by inflating LVF accuracy in the dyslexic

group. It should be noted however that in the current sample handedness showed no correspondence with individual variability in the presence of a RVF advantage, with all LVF advantage participants being right handed. Further work is needed to consider how handedness may interact with word recognition asymmetries in dyslexic and typical readers.

It should be noted that the experimental task used by Henderson et al. does deviate from standard design in visual half-field paradigms in a number of ways. Most salient is the difference in response measure and subsequent lack of a reaction time measure. Typically, visual half-field tasks require participants to verbally report the presented word, thus allowing for speed of naming to be measured as well as accuracy. In contrast, the current task required participants to type the word after presentation in a non-timed fashion. This lack of a response time measure seems somewhat counterintuitive with regards to the theory being tested, since one might assume that impaired callosal transfer might be most pronounced in a speed deficit. However, other studies of the bilateral advantage have reported this effect for both reaction time and accuracy measures (Mohr et al., 2007; Mohr & Pulvermüller, 2002). Further, a very short stimulus presentation of 60 ms was used, much shorter than the 150 ms presentation reported as standard by Hunter and Brysbaert (2008). Other studies with this paradigm have also used picture stimuli rather than written word stimuli, which could make the task easier for dyslexic participants (Hunter & Brysbaert, 2008). Given that this was a replication study, these design parameters were kept identical to the original study, however future work might consider systematically assessing the effect of such parameters.

6.4.4 *Future directions*

As outlined in the introduction, dominant models of word recognition predict that words presented in the LVF and so initially projected to the right hemisphere require callosal transfer for further processing (Ellis et al., 2009; Price & Devlin, 2011). Consequently, if we assume a deficit in this transfer as a characteristic of dyslexic reading, these models would predict a specific difficulty with reading visual word forms in the LVF and thus an increased RVF advantage, as outlined in Table 6.1.

What both these models and the interhemispheric deficit theory of dyslexia remain less clear on is the question of the nature of the representation that is or is not transferred across the corpus callosum. Is this simply a low level visual representation, or a more developed orthographical representation with encoding of letter information? Similarly, the interhemispheric deficit hypothesis of dyslexia needs clarification as to what aspect of interhemispheric communication is affected. The corpus callosum is a heterogeneous structure with distinct sub-regions that are involved in the transfer of different types of information at different levels of processing. For this theory to move forward therefore, it will be important for it to take into account what is known about the topography of the corpus callosum. Rather than assuming a global deficit in interhemispheric transfer, dyslexia could involve dysfunction in specific callosal segment(s) leading to disruption to the transfer of specific types of information. Specifying such details would enable more effective testing of the predictions of the theory, such that different aspects of interhemispheric transfer beyond that of visual information can be probed in dyslexia in the context of falsifiable predictions.

Further, it should be noted that both the original Henderson et al. study and the current study collected data from high functioning, highly compensated dyslexic students. As such, it is unclear to what extent differences in visual half field performance and underlying interhemispheric interactions may be the result of processes of compensation for reading difficulties, rather than involved with the onset of the disorder itself. Replication in samples of younger, more severe cases of dyslexia would be informative.

6.4.5 *Summary and conclusions*

This study highlights the importance of replication studies to ensure that a significant effect frequently cited in the literature is a true effect. It was found that while LVF performance in dyslexic participants does provide support for the interhemispheric deficit theory of dyslexia, the failure to replicate a reduced RBA in dyslexia does not. This pattern may however reflect the type of disruption to interhemispheric transfer in dyslexia, being more consistent with a model of degraded transfer of LVF word input to the left hemisphere than slow callosal transfer. This proposal should be further investigated by studies explicitly designed to test the predictions of the two models. Further work is also needed to refine the interhemispheric deficit theory of dyslexia, in order to generate more specific testable hypotheses as to which aspects of interhemispheric transfer are affected.

Chapter 7 Correspondence between Doppler ultrasound and the visual half-field method as measures of language lateralisation

Abstract

Comparing different methods of measuring asymmetries in brain function for language is of both practical and theoretical interest. The current study aimed to evaluate to what extent behavioural measures of laterality from the visual-half field (VHF) paradigm and imaging measures from functional transcranial Doppler ultrasound (fTCD) show agreement when used within the same individuals. 58 participants were administered a VHF task and the fTCD method with three language tasks. Overall, although modest levels of concordance were found between dominance classifications obtained from these methods, laterality indices demonstrated low and non-significant correlations. Consistency of lateralisation across the three fTCD tasks was found to be unrelated to VHF measures of interhemispheric transfer efficiency. However, there was preliminary evidence that the effect of strength of interhemispheric communication on language abilities depended on whether language processes shared a common lateralisation pattern; specifically, a negative relationship between the right visual field advantage and a measure of rapid naming ability was found in an inconsistently lateralised group only (although this did not survive correction for multiple comparisons). Overall, when comparing these different metrics of asymmetry, it is important to consider that they may be sensitive to different

aspects of interhemispheric interactions during language processing; they should thus be seen as complementary rather than conflicting.

7.1 Introduction

Multiple methods have been developed for assessment of cerebral lateralisation for language, including behavioural methods, brain imaging methods and the Wada technique for clinical cases. These methods each have their own associated advantages and disadvantages. One key consideration in assessing lateralisation is time and cost. Functional Magnetic Resonance Imaging (fMRI) offers a powerful tool for assessing lateralisation at a high spatial resolution; however, the high level of cost and unsuitability of certain individuals for this procedure (e.g. individuals with cardiac pacemakers and young children) present challenges for its widespread use. Two alternative methods that avoid these challenges are Doppler Ultrasound (fTCD) and the visual half-field technique. FTCD is an ultrasound method for measuring asymmetry of blood flow velocity in the left and right middle cerebral arteries. This procedure is suitable for use with young children, as the signal is relatively robust against head movements. FTCD offers a cheaper and more portable alternative to fMRI, but at the cost of regional specificity in laterality; FTCD LIs reflect asymmetry of blood flow supplied to a range of frontal, temporal and parietal areas covered by the middle cerebral artery territory.

The visual half-field (VHF) method offers a behavioural measure of cerebral asymmetry that can be easily administered to large groups. This method capitalises on a key feature of visual system architecture; the crossing of visual input at the optic chiasm. This results in representation of the left visual field (LVF) in right occipital cortex, and the right visual field (RVF) in left occipital cortex. The

VHF method relies on the interaction of this hemispheric separation of visual input at an early stage of visual processing, with hemispheric specialisation of later processing stages for visual word form input. That is, a word presented in the RVF will be projected to left visual cortex where it can then rely on within-hemisphere connections to access left lateralised specialised processing areas such as ventral occipitotemporal cortex and more anterior language areas. Conversely, a LVF presented word will be initially sent to right visual cortex, and therefore require transfer across the corpus callosum in order to access these left lateralised processing circuits. By briefly presenting visual word forms to the left and right of fixation for subsequent recognition, the VHF paradigm has demonstrated a RVF advantage for word recognition; compared to LVF words, RVF words are recognised faster and more accurately. This is attributed to the extra processing cost entailed in the need for transfer of this LVF input to the language dominant left hemisphere (Zaidel et al., 1990). Therefore, the VHF method has the potential to offer a cheaper and easier behavioural index of lateralisation for language.

7.1.1 *Comparisons between different measures of laterality*

Understanding how asymmetries reported from this range of methods relate to one another is of both practical and theoretical interest. At a practical level, it is important to know to what extent laterality measurements from these methods are interchangeable; if laterality classifications using behavioural methods show good agreement with those obtained from imaging methods, this would support the use of such behavioural methods in cases where administration of physiological techniques is impractical. A handfull of studies have compared VHF laterality measures with those obtained from imaging methods (Hunter &

Brysbaert, 2008; Krach et al., 2006; van der Haegen et al., 2011). While fMRI studies have reported a good level of correspondence, a study by Krach et al., (2006) reported poor correlations between VHF and fTCD laterality indices (r = .18), and only moderate concordance between their laterality classifications (74%). This was based on fTCD LIs calculated from word generation; although this study planned to compare laterality measures between the two methods using the same task, puzzlingly fTCD measures with the VHF lexical decision task revealed right lateralisation at the group level. This could reflect overshadowing of left lateralised word recognition processes by right lateralised directed attention employed during performance of VHF tasks. Overall, this argues against the interchangeability of fTCD and VHF measures of laterality.

At a theoretical level, these metrics of asymmetry may be sensitive to different properties of brain functioning that arise from lateralisation; comparison across methods could therefore offer insights into how these may manifest and interact. For example, intrinsic to the behavioural asymmetries observed with the VHF method is interhemispheric transfer; that is, the extent to which one is left lateralised for language affects the extent to which interhemispheric transfer of LVF input is required. However, this means that the classic metric of the RVF advantage is not only influenced by strength of laterality, but also by the efficiency of interhemispheric transfer. That is, a person may be strongly lateralised for language, and yet possess a high efficiency of interhemispheric transfer meaning that the additional cost associated with callosal transfer of LVF input is minimised, resulting in more balanced performance for RVF and LVF words. Conversely, another individual may possess the same level of language lateralisation but have poor efficiency of interhemispheric transfer, resulting in a larger processing cost

for LVF words and a stronger behavioural asymmetry. By contrast, fTCD measures of laterality are based on measures of blood flow supply in each hemisphere directly and do not capture dynamics of interhemispheric interactions. That is, two individuals who rely on the left hemisphere for language processing to the same extent, but who vary in their efficiency of interhemispheric transfer would presumably give similar measurements of laterality using fTCD methods. Thus, efficiency of interhemispheric transfer may be a confound that could lead to discrepant laterality measurements from fTCD and VHF techniques.

7.1.2 *The effect of inconsistent laterality on interhemispheric transfer*

In a previous chapter (*Chapter 5*) I used the fTCD technique to investigate an aspect of within-individual variability in lateralisation that may indirectly relate to interhemispheric communication. In that study, I measured lateralisation using fTCD for different language tasks within the same participants and investigated the effect of consistency in lateralisation across language functions within an individual on language abilities. This was with the hypothesis that being inconsistently lateralised across different language functions would represent a less efficient language network that would be associated with poorer language skills.

A further area of interest concerns the effects of such an inconsistent organisation for language on the efficiency of interhemispheric integration. One hypothesis could be that inconsistent laterality is a symptom of a disruption to interhemispheric interactions. The corpus callosum is considered a critical structure in the establishment of lateralised circuits during development, for example by mediating inhibition of the non-dominant by the dominant hemisphere (van der Knaap & van der Ham, 2011). Weaker interhemispheric communication

could therefore be seen to result in less successful establishment of hemispheric dominance for language, resulting in a more distributed network across the hemispheres. Therefore, this account would predict that inconsistent lateralisation on fTCD measures should be associated with weaker interhemispheric transfer on VHF measures (e.g. a greater RVF advantage).

An alternative possibility is that the development of inconsistent laterality for different language functions would place greater demands on interhemispheric communication and thus via use-dependent plasticity result in a strengthening of interhemispheric transfer. This account would thus predict that an inconsistent organisation for language would be associated with stronger interhemispheric transfer on VHF measures (e.g. a weaker RVF advantage).

7.1.3 *The effect of interhemispheric transfer efficiency on language abilities*

There has been increasing interest in the importance of interhemispheric connectivity for language ability. A relationship between diffusion tensor imaging (DTI) measures of corpus callosum structural integrity (i.e. fractional anisotropy or FA) and language/reading skills has been reported in both typically developing and impaired samples (Dougherty et al., 2007; Frye et al., 2008; Kim et al., 2006; Welcome & Joanisse, 2014). Further, in a study of individuals born pre-term, damage to the corpus callosum and anterior commissure was the most reliable predictor of presence of language impairment (Northam et al., 2012). However, the reported direction of the relationship between callosal FA and language/reading abilities has been inconsistent across studies; whilst some have reported a negative relationship (Dougherty et al., 2007; Frye et al., 2008),

others have found a positive relationship (Kim et al., 2006; Welcome & Joanisse, 2014).

One potential explanation for this inconsistency is that the effect of interhemispheric connectivity on language abilities may depend on the extent to which language functions are commonly lateralised. In my previous chapter (*Chapter 5*), I tested the hypothesis that having an inconsistent pattern of laterality for different language functions may increase risk for development of poorer language skills. This was based on the assumption that an inconsistently lateralised language network would place greater demands on interhemispheric integration, which is argued to be less efficient than communication within a hemisphere (Ringo et al., 1994). However, within this framework it is possible that a good level of interhemispheric transfer efficiency may buffer against the potential adverse effects of inconsistent lateralisation. In this case, language functions located in different hemispheres would still be able to communicate efficiently, leading to no observable effect on language skills. By contrast, a 'double hit' of inconsistent laterality and poor efficiency of interhemispheric transfer may be the necessary conditions for higher risk of impaired language abilities (see *Introduction: 1.5 Variability in lateralisation of different language sub-processes*). The current fTCD and VHF data allow for exploratory testing of this hypothesis in a preliminary (and small) sample. Specifically, this theory would predict that the efficiency of interhemispheric transfer as indexed by VHF measures should have no effect on language skills for consistently lateralised individuals; conversely in inconsistently lateralised individuals, reduced efficiency of interhemispheric transfer should be associated with poorer language abilities.

7.1.4 *Current study: Research Questions and predictions*

In this chapter, I combine measures of laterality obtained from fTCD and VHF methods within the same individuals to address the following research questions, and associated predictions:

Research question 1: Do laterality measurements demonstrate good concordance between fTCD and VHF dominance classifications, and a significant correlation between fTCD and VHF laterality indices?
Prediction: fTCD and VHF measures of laterality will demonstrate (i) a good level of concordance between dominance classifications and (ii) significant correlations between laterality indices.

Research question 2: Are there significant differences in VHF measures of interhemispheric transfer between consistently and inconsistently lateralised individuals as measured by fTCD?
Prediction: Inconsistently lateralised individuals will significantly differ from consistently lateralised individuals on VHF measures of interhemispheric transfer; this may manifest in either direction to indicate either significantly stronger or significantly weaker interhemispheric communication.

Research question 3: Is there evidence that having strong interhemispheric transfer can protect against the potential risk factor of inconsistent lateralisation for impaired language abilities?
Prediction: The size of the RVF advantage will be negatively correlated with scores on reading and language related measures in the inconsistently lateralised group only; there will be no significant relationship in the consistently lateralised group.

7.2 Methods

7.2.1 *Participants*

58 adult participants (17 male, 6 left handed) aged 18-45 took part in this experiment. Since data were collected as part of another study comparing VHF measures between dyslexics and controls (*Chapter 6*), 39 participants had a diagnosis of dyslexia and 19 had no diagnosis relating to reading or language impairment. Participants were recruited predominantly from local institutes of higher education. Exclusion criteria were significant hearing loss, history of neurological disease, head injury or epilepsy. All participants were native speakers of English.

7.2.2 *Materials and method*

7.2.2.1 Language assessment battery

Participants were tested with a comprehensive language assessment battery, details of which are outlined in a previous chapter (*Chapter 4: 4.2.3 Language assessment battery*). For the purposes of this study, data were used from two measures of reading/language skill: the Test of Word Reading Efficiency (TOWRE) and the York Adult Assessment battery Revised (YAA-R) rapid object naming sub-test. These measures were chosen due to their relevance to dyslexia, meaning the sample demonstrated a wide range of performance levels on these tests.

7.2.2.2 Visual half-field task

A computer-based visual half-field task was used for this study, adapted from Henderson, Barca and Ellis (2007). The procedure for this task has been described in a previous chapter (*Chapter 6: 6.2.2 Materials and method*).

Performance (accuracy) data from all three presentation conditions (LVF, RVF and BVF) were included in the current study.

A right visual field advantage score (RVF advantage) was calculated for each individual by subtracting their LVF accuracy score from their RVF accuracy score. A VHF laterality index (VHF LI) was then calculated by dividing this difference by the sum of LVF and RVF accuracy scores; this provides a metric of asymmetry from -1 (better LVF performance) to 1 (better RVF performance). A bilateral advantage score was calculated for each individual by subtracting their RVF accuracy from their BVF accuracy; this bilateral advantage score indexes the additional benefit of simultaneous presentation of redundant information to the other visual field. This measure is considered to reflect cooperative interactions between the hemispheres, and so can provide another VHF measure of the efficiency of interhemispheric communication. Again, to provide a normalised measure of relative differences between these conditions, a bilateral advantage index (BA index) was calculated by taking the difference between accuracy on BVF and RVF conditions and dividing this by the total accuracy across these conditions, (B-R)/(B+R).

7.2.2.3 fTCD tasks

Three tasks were used for laterality measurement with fTCD: phonological decision, semantic decision and sentence generation. The procedure for these tasks has been described in a previous chapter (*Chapter 5: 5.2.3 Measurement of lateralisation*). These were designed so as to follow a similar structure (illustrated in Fig. 7.1); following a 'CLEAR' cue for 3 seconds instructing participants to clear their mind, the language task was performed for 20 seconds,

followed by a rest period for 10 seconds. Each task consisted of 15 of these trials, lasting 33 seconds each.

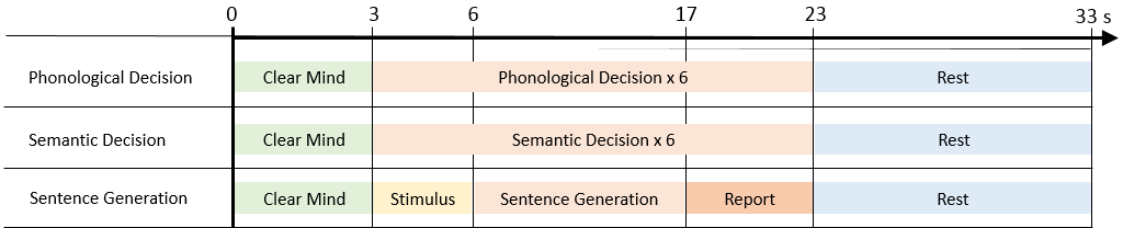


Figure 7.1: FTCD tasks. *Task design and timings of the three language tasks for use with fTCD.*

7.2.3 *Calculation of laterality indices and classification into dominance categories*

Laterality indices (LIs) were calculated for each participant for the VHF task and for each of the fTCD tasks individually, yielding a total of four LIs per participant. VHF LIs were calculated as outlined above using the formula (LVF-RVF)/(LVF+RVF). FTCD LIs were calculated using the mean method, as detailed in a previous chapter (*Chapter 5: 5.2.3.4 FTCD analysis and calculation of laterality indices*). Two methods were used to classify individuals into different dominance categories based on these LIs for each task. One method classified participants into left, right and bilateral groups using the cut-offs of 0.2 and -0.2 (with LI values between these cut-offs being classified as bilateral). A further method was used to classify participants into only typical versus atypical laterality groups, using zero as a cut-off in both methods (with a value of 0 being classed as atypical). For the fTCD data, these typical/atypical classifications were further used to categorise individuals into either consistent or inconsistent laterality groups, on the basis of whether their typical/atypical dominance classifications for all tasks agreed (consistent lateralisation) or whether one or more showed

discrepant laterality from the others (inconsistent lateralisation). In the current sample of 58 participants, 16 were inconsistently lateralised and 42 were consistently lateralised.

7.2.4 *Summary of measures*

Table 7.1 summarises the measures calculated from the VHF and fTCD methods for use in the statistical analyses below. This details how each measure is calculated and what it is interpreted to represent.

Table 7.1: Summary of measures. *An outline of key measures calculated from VHF and fTCD methods for use in statistical analyses (continued on next page).*

Measure name	Description of calculation	Interpretation
VHF measures		
RVF advantage	RVF accuracy minus LVF accuracy	Reflects both the strength of lateralisation for word recognition, and the efficiency of interhemispheric transfer. A higher value could reflect either stronger left lateralisation or weaker interhemispheric transfer.
VHF LI	(RVF - LVF)/(RVF + LVF)	As above, but gives relative difference between conditions rather than absolute difference and thus normalises for overall accuracy.
Bilateral advantage	BVF accuracy minus RVF accuracy	Reflects efficiency of communication between hemispheres during word recognition.
BA index	(BVF - RVF)/(BVF + RVF)	As above, but normalises for overall accuracy.

fTCD measures		
PD LI	Broadly calculated as the difference in blood flow velocity between the left and right middle cerebral arteries during the phonological decision task.	Reflects the strength of lateralisation of blood flow during phonological processing.
SD LI	Broadly calculated as the difference in blood flow velocity between the left and right middle cerebral arteries during the semantic decision task.	Reflects the strength of lateralisation of blood flow during semantic processing.
SG LI	Broadly calculated as the difference in blood flow velocity between the left and right middle cerebral arteries during the sentence generation task.	Reflects the strength of lateralisation of blood flow during covert speech production.
Consistency in lateralisation	Participants classified as consistent if all three fTCD LIs show the same dominance (LLL or RRR), or inconsistent if any of these disagree in their dominance.	Reflects the extent to which different language sub-processes share a common lateralisation pattern within an individual.

7.2.5 *Statistical analyses*

Research question 1: Do laterality measurements demonstrate good concordance between fTCD and VHF dominance classifications, and a significant correlation between fTCD and VHF laterality indices?

Concordance between dominance classifications obtained with fTCD and VHF methods was assessed for both methods of dominance classification (either with or without the category of bilaterality), and for all three fTCD tasks separately. Spearman's correlations were used to test the significance of correlations between laterality indices calculated from VHF and fTCD methods, again for each fTCD task separately.

Research question 2: Are there significant differences in VHF measures of interhemispheric transfer between consistently and inconsistently lateralised individuals as measured by fTCD?

To test for a difference in interhemispheric transfer between consistent (n = 42) and inconsistent (n = 16) laterality groups, a multi-level model analysis was run to compare RVF and LVF accuracy between groups. A significant interaction was predicted between visual field and group, in the form of a difference in the size of the RVF advantage between groups; this may manifest in either direction. The two groups were also compared on a further VHF metric of interhemispheric communication, the bilateral advantage, using a multi-level model analysis to compare RVF and BVF accuracy between groups. Again, a significant interaction was predicted between visual field and group, in the form of a difference in the size of the bilateral advantage between groups.

Research question 3: Is there evidence that having strong interhemispheric transfer can protect against the potential risk factor of inconsistent lateralisation for impaired language abilities?

To test the hypothesis that efficiency of interhemispheric transfer can buffer against the potential adverse effects of inconsistent language laterality on language abilities, I correlated the size of the RVF advantage with two measures of reading/language skills (TOWRE overall standard score and YAA-R rate of object naming) in consistent and inconsistent groups separately, with the prediction that a significant negative correlation would be found in the inconsistent group only. A Bonferroni corrected p value of .05/2 = .025 was used to correct for multiple comparisons.

7.3 Results

7.3.1 *Research question 1: Concordance and correlations between VHF and fTCD laterality measures*

Concordance between laterality classifications obtained from VHF and fTCD measures can be found in Table 7.2. As can be seen, concordance rates decrease dramatically when bilaterality is included as a third dominance category. The highest concordance between VHF and fTCD typical/atypical classifications is seen when sentence generation is used as the fTCD task. This is likely due to the fact that the majority of participants demonstrated left lateralisation for both VHF and sentence generation tasks.

Table 7.2: Method concordances. *Concordance between VHF and fTCD measures of laterality for (i) different classification methods (LRB = Left, right, bilateral) and (ii) different language tasks for fTCD laterality measurement (PD = Phonological decision, SD = semantic decision, SG = sentence generation).*

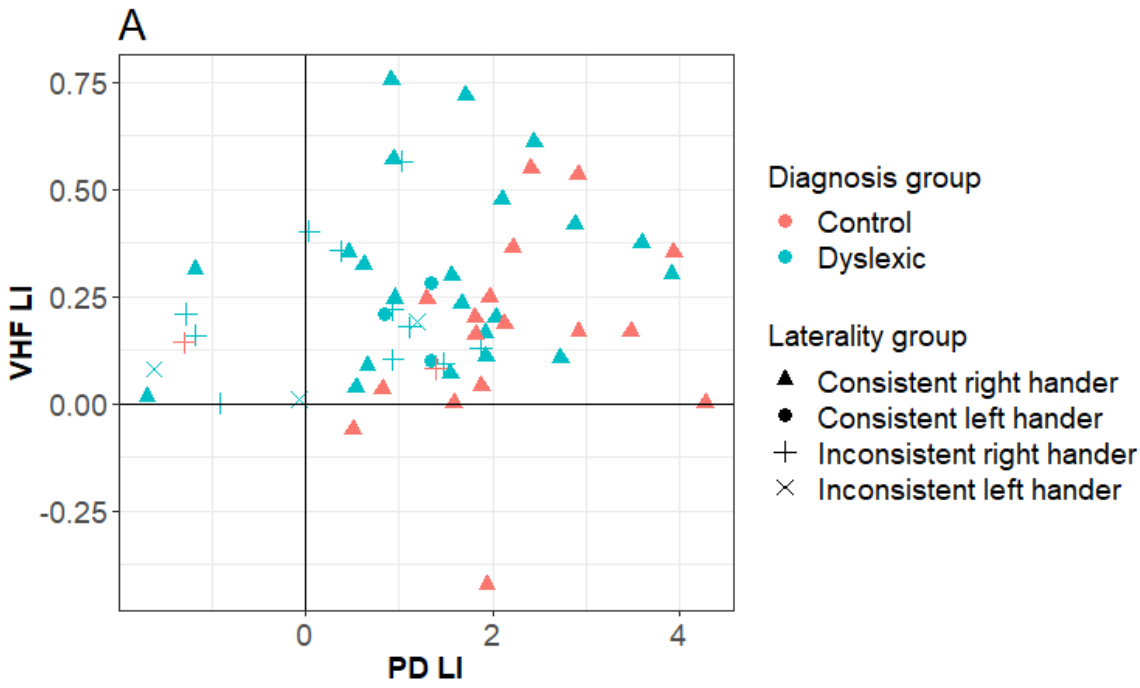
Task pair	LRB classification	Typical/atypical classification
VHF_PD	46.55%	81.03%
VHF_SD	41.37%	65.52%
VHF_SG	53.45%	89.66%

Correlations between laterality indices calculated from the two methods are given in Table 7.3, and shown in Figure 7.2. As can be seen, correlations are low and none reach statistical significance. Thus, despite the high concordance between VHF and sentence generation laterality measures, the correlation between LIs obtained using these methods is very low. Of particular note is the range of fTCD

LIs demonstrated by participants with a VHF LI of 0, many of whom are strongly lateralised on fTCD.

Table 7.3: Correlations between LIs. *Spearman's correlations between laterality indices calculated from the visual half-field method and laterality indices calculated from different language tasks with fTCD (PD = Phonological decision, SD = semantic decision, SG = sentence generation).*

Task pair	rho	p value
VHF_PD	0.22	0.097
VHF_SD	0.11	0.398
VHF_SG	0.07	0.596



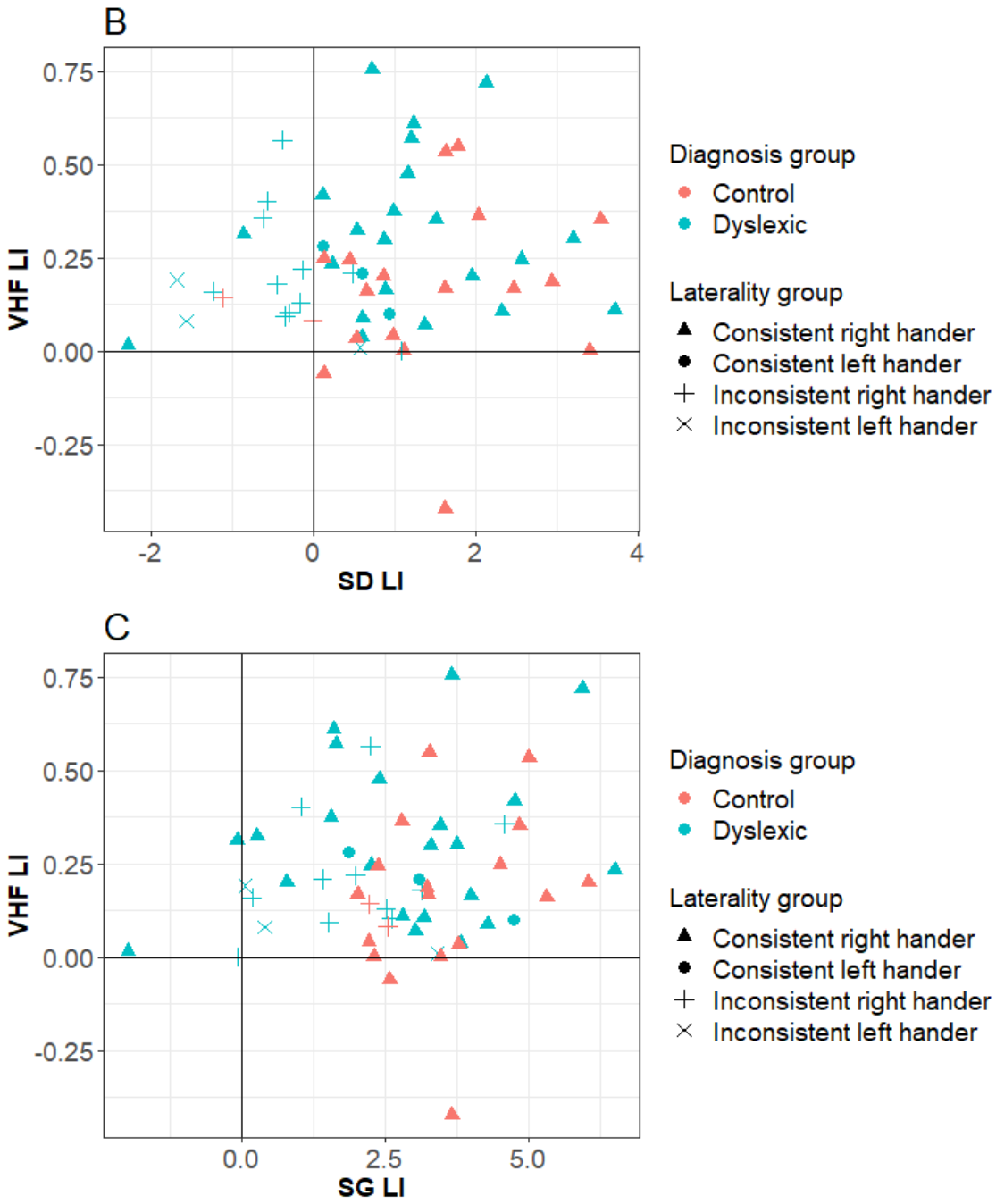


Figure 7.2: Relationships between VHF and fTCD laterality indices *on the phonological decision (A), semantic decision (B) and sentence generation (C) tasks. Control and dyslexic individuals are shown in different colours; shape codes for laterality group on fTCD measures (consistent versus inconsistent) and handedness.*

7.3.2 *Research question 2: Effect of lateralisation consistency on VHF measures of interhemispheric transfer*

To investigate whether inconsistent lateralisation across different language functions was associated with altered levels of interhemispheric transfer efficiency, a multi-level modelling analysis was run to compare accuracy on LVF and RVF conditions between consistent and inconsistent laterality groups. Accuracy of these two groups on the two conditions is shown in Figure 7.3. To test for a significant interaction between condition and group, two random intercept models were compared; a null model with a random effect of participant and fixed effects of group and condition (with no interaction), and a full model with the same random effect of participant but an interaction between group and condition. A likelihood ratio test indicated that the full model did not provide a better fit to the data, ($\chi^2(1) = 0.203$, $p = 0.652$). Therefore, there was no significant interaction between group and condition, indicating that the inconsistent group did not significantly differ from the inconsistent group on the RVF advantage.

To test whether the two groups varied on another VHF measure of interhemispheric communication, the bilateral advantage, the same multi-level model analysis was conducted using the two conditions of RVF and BVF. Mean values for each group along with individual data points for each condition are given in Figure 7.4. Again, a likelihood ratio test indicated that a full model with an interaction between group and condition did not provide a better fit to the data than the null model with no interaction ($\chi^2(1) = 0.079$, $p = 0.779$).

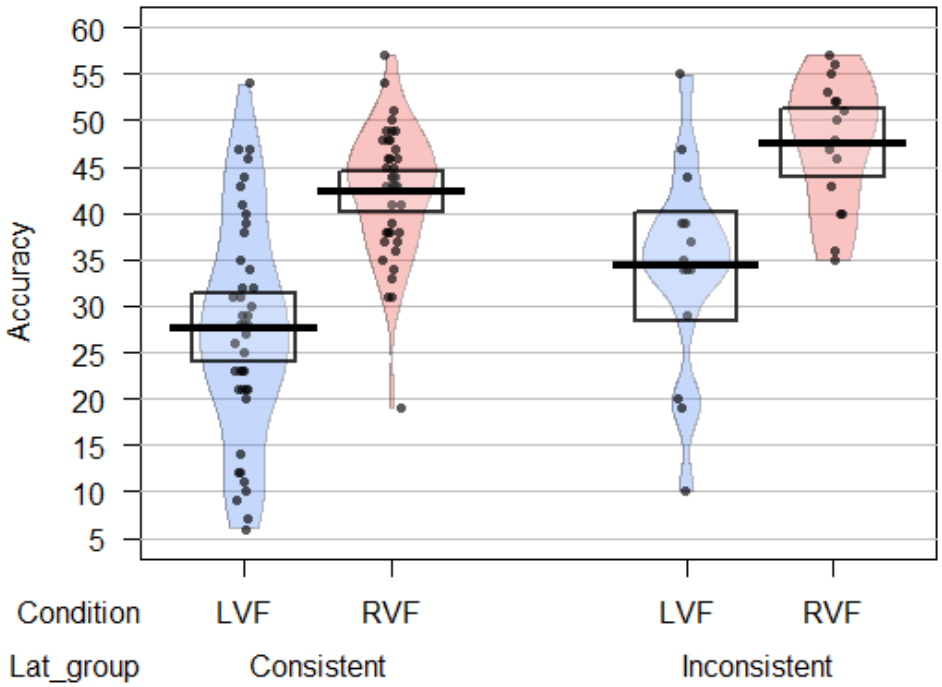


Figure 7.3: RVF advantage in laterality groups. *Accuracy (out of 60) on the LVF and RVF conditions of the VHF task in the consistent and inconsistent laterality groups. Bars indicate group means, boxes indicate 95% confidence intervals.*

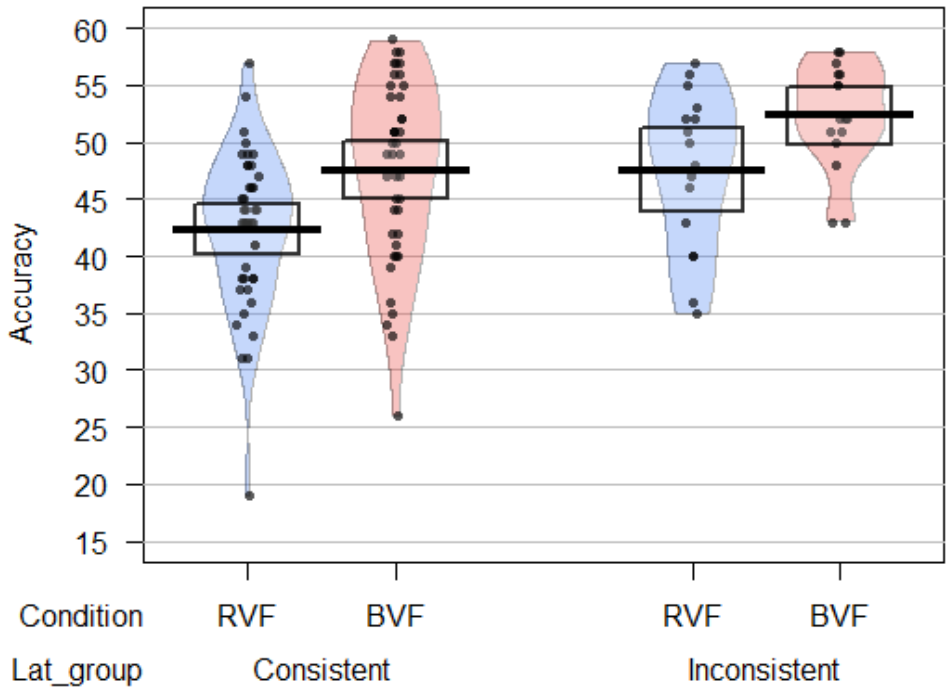


Figure 7.4: Bilateral advantage in laterality groups. *Accuracy (out of 60) on the RVF and BVF conditions of the VHF task in the consistent and inconsistent laterality groups. Bars indicate group means, boxes indicate 95% confidence intervals.*

7.3.3 *Research question 3: Interhemispheric transfer efficiency as a protector against inconsistent laterality*

To explore whether a high level of interhemispheric transfer efficiency may protect against the potential risk factor of inconsistent laterality on language abilities, RVF advantage scores were plotted against scores on standardised reading assessments for inconsistent and consistent groups. This was done for two measures of reading/language ability; the TOWRE (test of word reading efficiency) and YAA-R object naming (York Adult Assessment battery Revised). Figure 7.5 illustrates the relationships seen for TOWRE scores; correlation analyses found no significant relationship with RVF advantage scores in either the consistent (r = -0.140, p = .375) or the inconsistent group (r = -0.472, p = .065). However, when the same analyses were done for YAA-R object naming rate scores, a significant negative correlation was found in the inconsistent group (r = -0.542, p = .030) but not the consistent group (r = -0.040, p = .799). These relationships are illustrated in Figure 7.6. The same pattern of results was found when using VHF LIs instead of RVF advantage scores. This suggests that in this small sample of individuals with inconsistent laterality, having a smaller RVF advantage (indicative of more efficient interhemispheric transfer) is associated with improved rapid naming performance. However, this significant result did not survive the Bonferroni correction for multiple comparisons using a corrected p value of .025.

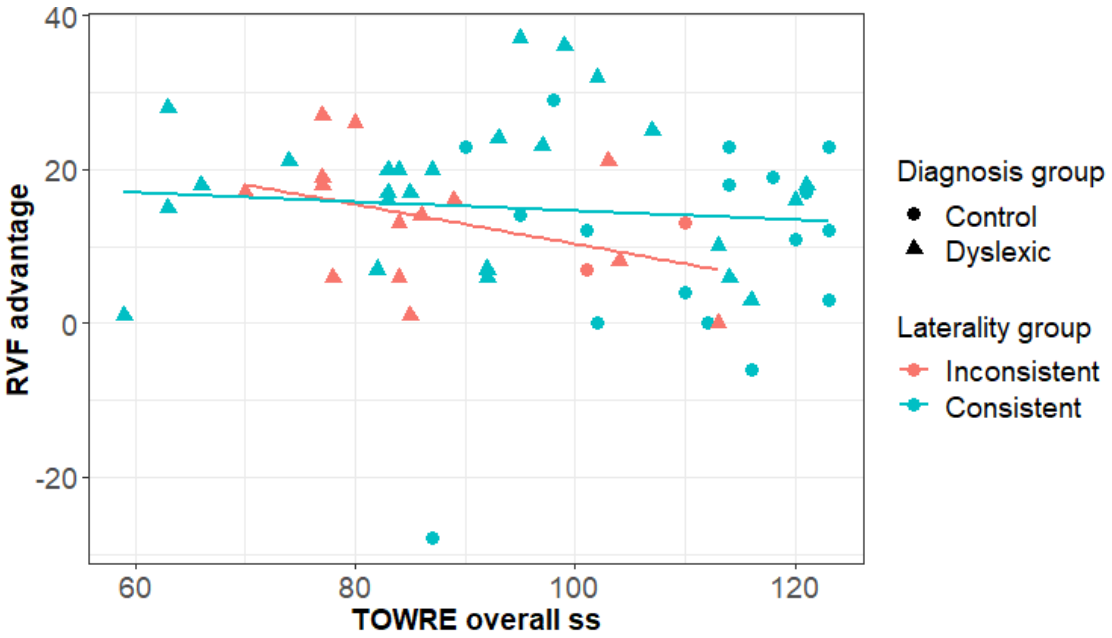


Figure 7.5: RVF advantage and TOWRE scores. *Scatter plot shows the relationship between RVF advantage scores and TOWRE overall standard scores for inconsistent and consistent laterality groups; regression lines are given for each group separately.*

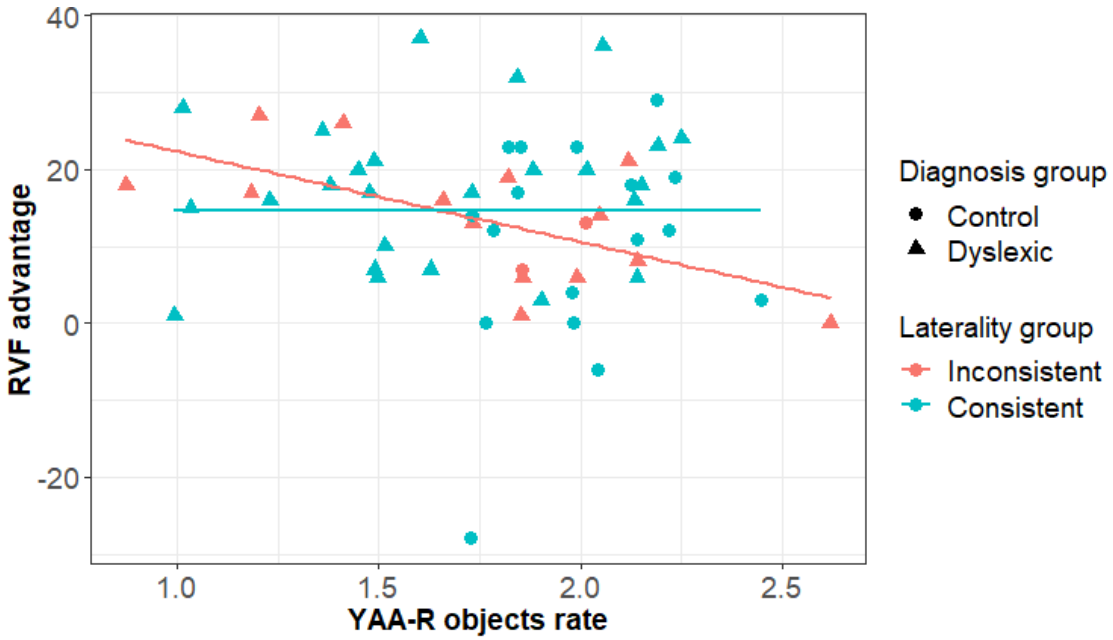


Figure 7.6: RVF advantage and YAA-R rapid naming scores. *Scatter plot shows the relationship between RVF advantage scores and YAA-R object rapid naming rate scores for inconsistent and consistent laterality groups; regression lines are given for each group separately.*

7.4 Discussion

7.4.1 *Summary of findings*

This chapter aimed to combine and compare VHF and fTCD measures of laterality in order to address a range of practical and theoretical questions regarding their relationship to one another. Overall, agreement between laterality measurements obtained from these two methods was poor, suggesting they should not be viewed as interchangeable. The hypothesised differences in VHF measures of interhemispheric communication between participants with consistent and inconsistent lateralisation as measured using fTCD were not found. When considering relationships between measures of reading/language skills and VHF measures of interhemispheric transfer efficiency, there was preliminary evidence that in some cases this depends on consistency of lateralisation; for a measure of rapid naming, a significant negative relationship was found for the inconsistent group only (no significant relationship was found in the consistent group). However, this result did not survive correction for multiple comparisons and was not replicated for a measure of word and non-word reading (TOWRE).

7.4.2 *Correspondence between VHF and fTCD measures of laterality*

Concordances between dominance classifications obtained from VHF and fTCD measures of laterality depended greatly on whether bilaterality was included as a category. When included, concordances with VHF dominance were low for all three fTCD tasks, ranging from around 40-50%. This is consistent with previous reports that classifications of bilaterality are less reliable within subjects across different methods of measurement (Jansen et al., 2006). Using a simple cut-off of 0 to classify participants into typical versus atypical categories yielded higher

levels of concordance across the two methods, ranging from around 65-90%. The highest level of concordance was observed when using the sentence generation task for fTCD measurement. However, this high concordance is likely simply attributable to the fact that the majority of the sample were left lateralised for both VHF and SG tasks. This more trivial explanation for the high level of agreement between these methods is supported by the low correlations found between their laterality indices. Correlations with VHF LIs across all three fTCD task LIs were low and not statistically significant. The findings from this analysis thus agree with previous work reporting an overall poor level of correspondence between laterality measurements from VHF and fTCD methods (Krach et al., 2006).

There are multiple possible explanations for this poor level of correspondence between these imaging and behavioural measures of laterality. Firstly, it is possible that the behavioural VHF method is simply a poor indicator of underlying asymmetries in brain function. Whilst it can yield robust effects at the group level, with the RVF advantage being highly reproducible across samples (Brederoo et al., 2019), it may not be an appropriate method for accurately measuring individual variability in lateralisation of underlying brain areas. Indeed, the validity of the VHF method as a measure of cerebral lateralisation has long been questioned, with arguments that these behavioural asymmetries may instead reflect biases in attention allocation across the visual field (Heron, 1957).

Alternatively, it is possible that the fTCD and VHF methods are measuring different (but equally valid) aspects of interhemispheric dynamics, that together provide a broader picture of the nature of cerebral lateralisation for language processing within an individual. As previously discussed, VHF LIs will be affected not only by the degree to which word recognition processes are lateralised, but

further by the level of efficiency of interhemispheric transfer of visual word input from the non-dominant to the dominant hemisphere. For example, in the current study there were a number of participants with a VHF LI of 0. Rather than indicating use of both hemispheres for word recognition to an identical extent, this could instead reflect a level of efficiency of interhemispheric transfer that was sufficient to result in no additional processing cost for LVF compared to RVF words, despite left lateralisation of word recognition. This interpretation is supported by fMRI evidence from van der Haegen, Cai, Seurinck and Brysbaert (2011), who reported that participants who failed to show a clear VHF asymmetry did not demonstrate reduced laterality on fMRI, instead showing clear dominance to one hemisphere. Indeed in the current study, most participants with a VHF LI of zero nevertheless demonstrated strong laterality for the fTCD tasks. This suggests that caution should be exercised when trying to draw inferences on underlying lateralisation from an observed absence of an asymmetry on VHF measures.

A final possible interpretation of the current lack of correspondence between methods is that the current VHF task was not optimally designed. Previous failures to find good agreement between VHF and imaging measures of laterality were criticised by Hunter and Brysbaert (2008) as being attributable to the use of sub-optimal VHF paradigms. In particular, they highlighted the importance of including a sufficient number of trials, using bilaterally balanced presentation, controlling fixation location, and employing a task that requires an overt naming response. Although the current study improved on the design of the VHF task used by Krach et al., (2006) by substantially increasing the number of trials, it does not meet all recommended requirements. In particular, the paradigm used

here differs from standard VHF tasks in that participants were required to respond by typing in the word using a keyboard, rather than via verbal report. The design of the task was constrained since this data was collected in the context of a replication study; task design was thus kept identical to the original study except for the inclusion of additional trials (L. Henderson et al., 2007). However, lateralisation of speech generation may yield more robust measurements (Bradshaw, Thompson, Wilson, Bishop, & Woodhead, 2017), and in the case of this study would be more similar to the processes engaged by the fTCD tasks. Thus it is possible that greater correspondence between fTCD and VHF measures of laterality may be possible if VHF task design met these requirements, as has been demonstrated for fMRI (Hunter & Brysbaert, 2008; van der Haegen et al., 2011).

7.4.3 *Effect of inconsistency in lateralisation on interhemispheric transfer*

When comparing consistent and inconsistent laterality groups (as defined by fTCD measurements across multiple language tasks), the hypothesised differences on VHF measures of interhemispheric transfer were not found. These groups demonstrated no difference in either the RVF advantage or the bilateral advantage. This does not support the idea that changes in interhemispheric connectivity have downstream consequences for development of lateralisation patterns for language. It should be noted however that more direct measures of structural integrity of interhemispheric tracts have yielded supporting evidence; a recent study of a large fMRI meta-analytic dataset reported that brain regions that were more strongly lateralised demonstrated lower connectivity within the corpus callosum on micro- and macro-structural measures (Karolis et al., 2019). The RVF and bilateral advantages used in the current analyses may not be sensitive

enough measures of interhemispheric connectivity to reveal such a relationship. In particular, it may be that a high level of regional specificity is needed, both for measures of lateralisation and measures of interhemispheric connectivity, in order to reveal regionally specific relationships as demonstrated by Karolis et al., (2019).

7.4.4 *Preliminary test of the double-hit hypothesis*

The combining of fTCD laterality and VHF interhemispheric transfer measures in the context of this thesis allowed for a preliminary test of the 'double-hit hypothesis' of language impairment. In this hypothesis, I proposed that a 'double hit' of inconsistent laterality and poor efficiency of interhemispheric transfer are the necessary conditions for increasing risk of impaired language abilities. The current analysis found tentative evidence in support of this hypothesis for a measure of rapid naming. Specifically, a significant relationship between the RVF advantage and rapid naming skill was found for the inconsistent laterality group only, with no relationship in the consistent group. This is what one would expect if it is assumed that changes in interhemispheric connectivity should only have an impact on language functioning if different language processes are lateralised to different hemispheres. Conversely, when all processes share a common lateralisation to one hemisphere, demands on interhemispheric communication for language functioning are greatly reduced, resulting in no observed relationship.

However, although a trend towards this pattern was seen for an alternative measure of reading (TOWRE), relationships with the RVF advantage were non-significant for both laterality groups using this measure. Furthermore, the significant correlation reported for the inconsistent group using the YAA-R rapid

naming measure did not survive correction for multiple comparisons. Overall, this analysis should be seen as highly exploratory and limited by the small sample size (with only 16 inconsistently lateralised participants). It demonstrates the type of pattern one would expect to see if the double hit hypothesis were correct. For a more conclusive test, this type of analysis should be replicated in a larger sample with an *a priori* power analysis, using more direct measures of interhemispheric transfer efficiency such as corpus callosum FA or ERP measures of interhemispheric transfer time (e.g. Friedrich et al., 2017).

7.4.5 *Limitations of the study*

Some limitations of the study have already been highlighted in the above discussions. Power is an area of concern; due to the small sample sizes and lack of an *a priori* power analysis, it is likely that many of the analyses were underpowered. In particular, the small number of participants classed as inconsistent that could be included in the current study (i.e. that had given data for both VHF and fTCD tasks) makes the comparisons between consistent and inconsistent groups presented here highly exploratory and in need of replication. Further, the small number of participants with atypical laterality (especially on the VHF task) is a limitation; as discussed, this likely resulted in an artificially high concordance between VHF LIs and sentence generation fTCD LIs (borne out by the very low correlation between them). Also of significance is the sub-optimal design of the VHF task as discussed above, which may have led to lower agreement between the methods of laterality measurement. Future studies comparing VHF LIs with those obtained from imaging methods should ensure robust VHF task design, as outlined by Hunter and Brysbaert (2008). Finally, it could be argued that the high number of dyslexic individuals within the current

sample may have affected results by potentially weakening the strength of correlations between methods. However, given the results were in agreement with a previous study reporting a poor level of fTCD-VHF correspondence in a sample of typically developing individuals (Krach et al., 2006), the low correlations observed do not seem to be specific to samples including reading disordered individuals. The inclusion of these participants provided a dataset with a good level of variability in standardised measures of reading, which was important for the analyses of relationships between these scores and VHF measures of interhemispheric transfer efficiency.

7.4.6 *Summary and conclusions*

In bringing together two very diverse measures of lateralisation for language, this chapter has highlighted some of the issues involved in interpreting different metrics of such asymmetries. Of particular importance is the need to appreciate that VHF measures of laterality likely also incorporate efficiency of interhemispheric transfer. This may explain the current failure to find a good level of agreement between this behavioural measure and a brain imaging measure of laterality. In particular, an absence of asymmetry on the VHF paradigm may in many cases be better understood as indicating greater efficiency of interhemispheric transfer, rather than a bilateral organisation for language. Future work aiming to validate the VHF method against brain imaging measures should ensure optimal design of the VHF task, which may increase robustness of results and thus concordance with imaging measures (Hunter & Brysbaert, 2008). The current analysis did find preliminary evidence in support of the double hit hypothesis of language impairment, although this is limited by the small sample and the use of an indirect measure of interhemispheric transfer efficiency. This

hypothesis should be tested more convincingly in a better powered sample using a more direct measure of interhemispheric connectivity such as those offered by DTI. Overall, different methods of laterality measurement should be seen as complementary, and when used together may offer a more holistic picture of an individual's pattern of interhemispheric interactions during language processing.

Chapter 8 Overall Discussion

This thesis has taken a novel approach to answering a long standing question in the field of the neuroscience of language. In so doing, it has aimed to bring research on the relationship between laterality and language impairment up to date with modern thinking on the mechanisms that give rise to cerebral asymmetries. This has involved considering the multidimensionality of lateralisation, as well as the importance of interhemispheric interactions in language functioning.

8.1 Language lateralisation and language impairment

Overall, the evidence reviewed and presented in this thesis suggests that reduced laterality is not associated with impaired language skills. Significant results reported in previous literature appear highly likely to be false positives attributable to underpowered samples. When larger scale studies are done, such as that by Wilson and Bishop (2018) with children with DLD and the current study of adults with developmental disorders (*Chapter 5*), no evidence for weaker strength of laterality is found. This suggests that research and theorising should stop pursuing this line of thought in the search to identify how the language system might be disrupted in language impairment.

However, the current findings suggest that aspects of lateralisation other than strength may be important. In Chapter five, an fTCD study found that a sample of adults with developmental disorders were more likely to demonstrate an inconsistent pattern of lateralisation across different language tasks than controls. This was also reflected in weaker correlations between task laterality indices, and increased variability in the direction of laterality across trials within

performance of a phonological task. These patterns resonate with the higher variability in lateralised responses reported by Kohler et al., (2015) in pre-school children. In this age group, such increased variability was associated with better language functioning, interpreted as increased plasticity. One hypothesis is that adults with developmental disorders retain this more infant-like pattern of laterality in which lateralised responses are more variable, both across different language functions and during performance of a single language task. Whilst early on in language development this higher plasticity in the lateralised response may be beneficial for language development, its persistence into adulthood may not be. This highlights the importance of considering the possible interaction between age and laterality; by taking a single 'snap-shot' measurement in early adulthood we may well be missing an earlier critical period in which differences between those on a trajectory towards impaired language and those with normal language development are most evident. While laterality changes later in life are well-established (Cabeza, 2002), changes in laterality at the early stages of development are not well understood, but may provide key insights into what puts some children at risk for language disorders.

However, whilst Chapter four demonstrated that this sample of 'high functioning' adults with developmental disorders did experience impaired language abilities, inconsistent laterality across the whole sample was not associated with poorer language functioning. It is possible therefore that this pattern of increased variability in the lateralised response may not be a cause of language impairment, but instead may be a consequence of experience of a developmental disorder. Given the unusual nature of the current cohort, being highly compensated and high achieving university students with developmental disorders, it is even

possible that such laterality differences could be the result of compensatory processes. Overall, the findings suggest the need to reconsider how we conceptualise the directionality of the relationship between altered laterality and language impairment.

8.2 Interhemispheric processing and dyslexia

My third study (*Chapter 6*) found evidence of a stronger right visual field (RVF) advantage on the visual half-field (VHF) paradigm in right handed adults with dyslexia compared to controls. The poor correspondence between laterality indices from VHF and fTCD methods found in Chapter 7 suggests that one should not interpret this as evidence that dyslexic individuals have stronger lateralisation for word recognition than controls. Instead, a more plausible interpretation is that this indicates weaker interhemispheric transfer of left visual field word form input from the right to the left (dominant) hemisphere. Such a finding fits well with previous reports of functional and structural abnormalities of interhemispheric connectivity in dyslexia (Daini et al., 2018; Paul, 2011; Welcome & Joanisse, 2014).

Future work should seek to directly relate functional and structural measures of interhemispheric transfer to each other in dyslexic samples. It has already been established in typically developing samples that ERP estimates of interhemispheric transfer times (IHTTs) are negatively correlated with fractional anisotropy (FA) within the corpus callosum (Westerhausen, Kreuder, Woerner, et al., 2006). Further, there has been preliminary work attempting to relate changes in IHTTs and DTI measures of callosal structural integrity to each other within clinical samples, such as schizophrenia and alcoholism (Schulte, Sullivan, Müller-Oehring, Adalsteinsson, & Pfefferbaum, 2005; Whitford et al., 2011). I would

predict that combining ERP and DTI measures in dyslexic participants would reveal significantly longer IHTTs that would be associated with significantly reduced corpus callosum FA. Such a demonstration of changes in interhemispheric structural connectivity that have ramifications for the functional dynamics of interhemispheric transfer would provide strong evidence for the interhemispheric deficit hypothesis of dyslexia.

8.3 The double-hit hypothesis of language impairment

This thesis has outlined a novel hypothesis concerning the neural factors that may increase risk for language impairments. Specifically, I have proposed that greater risk for impaired language skills may result from a 'double-hit' of (i) increased variability in the side of lateralisation for different language functions and (ii) weakened interhemispheric communication due to disrupted white matter connectivity in interhemispheric tracts. This would result in a failure of adequate or efficient communication between language sub-processes, causing impairments in language functioning.

The experimental work presented in this thesis does provide some consistent evidence with this hypothesis. Individuals with development disorders affecting language and literacy were found to demonstrate both an increased likelihood of an inconsistent pattern of lateralisation across different language functions (*Chapter 5*) and weaker interhemispheric transfer of visual word form information from the non-dominant to the dominant hemisphere (*Chapter 6*). As already highlighted, an inconsistent pattern of laterality for different language processes was not by itself found to be associated with poorer language skills across a sample of both impaired and typically developing adults. However, this hypothesis would predict that an absence of this effect could be attributable to

individual differences in strength of interhemispheric connectivity. That is, an inconsistent pattern of laterality would not have detrimental consequences on language if interhemispheric communication was sufficiently efficient.

I was able to conduct a very preliminary test of this proposal in the current thesis by combining fTCD measures of lateralisation and VHF measures of interhemispheric transfer within the same individuals (*Chapter 7*). The pattern predicted by this hypothesis was found for one measure of language skills; for rapid naming, a significant negative relationship was found between the RVF advantage (a proxy of interhemispheric transfer efficiency) and performance for an inconsistently lateralised group only, with no such relationship in the consistently lateralised group. Although consistent with the predictions of the double-hit hypothesis, this evidence is not compelling for numerous reasons; the analysis was highly exploratory and likely underpowered, and the result did not survive correction for multiple comparisons or replicate for another measure of word and non-word reading ability. This result is presented in the current thesis simply as an illustration of the type of pattern this hypothesis would predict. More robust testing of this hypothesis requires the integration of functional measures of lateralisation and direct structural measures of interhemispheric connectivity (e.g. DTI) within the same individuals. Overall, this double-hit hypothesis seeks to direct research towards considering the importance of the global interplay between the hemispheres during language functioning, rather than simply focusing on isolated asymmetries of functional activation.

8.4 Measurement of lateralisation with different methods

In my final chapter (*Chapter 7*), I reported poor correlations between laterality indices obtained within the same individuals using the VHF and fTCD methods.

This raises questions as to how we should react when different measures of laterality don't agree within individuals. I would argue that such lack of agreement does not necessarily indicate superiority of one measure over the other. Instead, it is important to recognise that these different methods are sensitive to different aspects of the dynamics of interhemispheric interactions. The fact that the VHF method yields a significant difference between dyslexic and control groups that would be predicted from the interhemispheric deficit hypothesis suggests that it is measuring something meaningful. In conjunction, the fTCD and VHF results suggest that these individuals may possess typical strength of lateralisation for language, but abnormal interhemispheric transfer of visual word form information that results in a larger behavioural asymmetry. In this way, different methods can be seen as complementary, and when used together can provide a more holistic picture of an individual's pattern of interhemispheric communication.

On the other hand, the systematic review of protocols used in the fMRI laterality literature presented in Chapter two of this thesis highlights the importance of standardisation across the field in the use of a particular method. This review revealed the high level of variability in the protocols used by fMRI laterality studies for quantification and classification of language laterality, and synthesised evidence demonstrating the effects these methodological differences can have on the resulting laterality measurements. In the face of such methodological variability, meta-analytic review of the literature is prevented, and it becomes difficult to integrate findings across studies. Similar issues exist for the fTCD method; in particular, Chapter 5 highlights the differences between the traditional 'peak' method of LI calculation and the newer 'mean' method. The advantages of the latter in terms of a more normal distribution of LIs suggest that this method

should be adopted in fTCD laterality research going forward. Overall, striving towards convergence in the field on robust standardised methodological protocols will be pivotal for progress in our theoretical understanding of language lateralisation.

Furthermore, the systematic review of different tasks used in fMRI laterality research presented in Chapter three highlights the variability in laterality that can be obtained with different tasks. Some of this variability is clearly attributable to arbitrary differences in protocols such as the use of different baseline tasks and lack of optimisation of the region of interest chosen. However, in order to properly investigate differences in extent of lateralisation for different language functions, a systematic approach must be taken in which laterality task batteries are employed; such tasks should be rigorously matched for non-linguistic parameters, but engage distinct aspects of language processing. The main experiment of the current thesis (*Chapter 5*) aimed to provide one such battery, in particular by developing a set of highly matched semantic/phonological decision tasks. This enabled contrasts between laterality obtained with each task both at a group level and within individuals, in order to demonstrate that semantic and phonological processing can dissociate in their patterns of laterality. Indeed, this was the most frequent pattern of inconsistent lateralisation found in this study. This demonstrates how such an approach can provide more conclusive evidence on differences in extent of lateralisation of different language functions, which can then be used to inform neurocognitive models of speech and language.

8.5 Concluding remarks

Overall, this thesis highlights the importance of considering within-individual variability in language lateralisation. The comparison of lateralisation across

different language functions suggests that in some cases this can dissociate within individuals. This supports a multidimensional over a unitary view of language lateralisation. By considering such multidimensionality, this thesis has furthered our understanding of how patterns of laterality may differ in developmental disorders of language. Answers to the question of whether lateralisation is beneficial for language functioning remain elusive; however, light may be shed by newer approaches to research that consider the mechanisms underlying asymmetry of functional activation, such as the dynamics of interhemispheric interactions.

Bibliography

Abbott, D. F., Waites, A. B., Lillywhite, L. M., & Jackson, G. D. (2010). fMRI assessment of language lateralization: An objective approach. *NeuroImage*, *50*(4), 1446–1455. https://doi.org/10.1016/j.neuroimage.2010.01.059

Aboitiz, F., López, J., & Montiel, J. (2003). Long distance communication in the human brain: Timing constraints for inter-hemispheric synchrony and the origin of brain lateralization. *Biological Research*, *36*(1), 89–99. https://doi.org/10.4067/S0716-97602003000100007

Adcock, J. E., Wise, R. G., Oxbury, J. M., Oxbury, S. M., & Matthews, P. M. (2003). Quantitative fMRI assessment of the differences in lateralization of language-related brain activation in patients with temporal lobe epilepsy. *NeuroImage*, *18*(2), 423–438. https://doi.org/10.1016/S1053-8119(02)00013-7

Allendorfer, J. B., Hernando, K. A., Hossain, S., Nenert, R., Holland, S. K., & Szaflarski, J. P. (2016). Arcuate Fasciculus Asymmetry Has a Hand in Language Function but not Handedness. *Human Brain Mapping*, *37*(9), 3297–3309. https://doi.org/10.1002/hbm.23241

Altmann, L. J. P., Lombardino, L. J., & Puranik, C. (2008). Sentence production in students with dyslexia. *International Journal of Language & Communication Disorders*, *43*(1), 55–76. https://doi.org/10.1080/13682820701284522

Annett, M. (1985). *Left, right, hand and brain: The right shift theory*. London: Lawrence Erlbaum.

Archibald, L. M. D., Cardy, J. O., Ansari, D., Olino, T., & Joanisse, M. F. (2019). The consistency and cognitive predictors of children's oral language, reading, and math learning profiles. *Learning and Individual Differences*, *70*, 130–141. https://doi.org/10.1016/j.lindif.2019.02.003

Archibald, L. M. D., Cardy, J. O., Joanisse, M. F., & Ansari, D. (2013). Language, Reading, and Math Learning Profiles in an Epidemiological Sample of School Age Children. *PLoS ONE*, *8*(10). https://doi.org/10.1371/journal.pone.0077463

Astle, D. E., Bathelt, J., & Holmes, J. (2019). Remapping the cognitive and neural profiles of children who struggle at school. *Developmental Science*, *22*(e12747). https://doi.org/10.1111/desc.12747

Baciu, M., Juphard, A., Cousin, E., & Bas, J. F. Le. (2005). Evaluating fMRI methods for assessing hemispheric language dominance in healthy subjects. *European Journal of Radiology*, *55*(2), 209–218. https://doi.org/10.1016/j.ejrad.2004.11.004

Baciu, M. V, Watson, J. M., Maccotta, L., McDermott, K. B., Buckner, R. L., Gilliam, F. G., & Ojemann, J. G. (2005). Evaluating functional MRI procedures for assessing hemispheric language dominance in

neurosurgical patients. *Neuroradiology*, *47*(11), 835–844. https://doi.org/10.1007/s00234-005-1431-3

Backes, W. H., Deblaere, K., Vonck, K., Kessels, A. G., Boon, P., Hofman, P., … Aldenkamp, A. P. (2005). Language activation distributions revealed by fMRI in post-operative epilepsy patients: Differences between left- and right-sided resections. *Epilepsy Research*, *66*(1–3), 1–12. https://doi.org/10.1016/j.eplepsyres.2005.06.007

Badcock, N. A., Bishop, D. V. M., Hardiman, M. J., Barry, J. G., & Watkins, K. E. (2012). Co-localisation of abnormal brain structure and function in specific language impairment. *Brain and Language*, *120*(3), 310–320. https://doi.org/10.1016/J.BANDL.2011.10.006

Badcock, N. A., & Groen, M. A. (2017). What can functional Transcranial Doppler Ultrasonography tell us about spoken language understanding? *Language, Cognition and Neuroscience*, *32*(7), 818–828. https://doi.org/10.1080/23273798.2016.1276608

Badcock, N. A., Nye, A., & Bishop, D. V. M. (2012). Using functional transcranial Doppler ultrasonography to assess language lateralisation: Influence of task and difficulty level. *Laterality: Asymmetries of Body, Brain and Cognition*, *17*(6), 694–710. https://doi.org/10.1080/1357650X.2011.615128

Badzakova-Trajkov, G., Hamm, J. P., & Waldie, K. E. (2005). The effects of redundant stimuli on visuospatial processing in developmental dyslexia. *Neuropsychologia*, *43*(3), 473–478. https://doi.org/10.1016/j.neuropsychologia.2004.06.016

Bailey, C. E., Manis, F. R., Pedersen, W. C., & Seidenberg, M. S. (2004). Variation among developmental dyslexics: Evidence from a printed-word-learning task. *Journal of Experimental Child Psychology*, *87*(2), 125–154. https://doi.org/10.1016/j.jecp.2003.10.004

Banich, M. T., & Brown, W. S. (2000). A life-span perspective on interaction between the cerebral hemispheres. *Developmental Neuropsychology*, *18*(1), 1–10. https://doi.org/10.1207/S15326942DN1801_1

Barca, L., Cornelissen, P., Simpson, M., Urooj, U., Woods, W., & Ellis, A. W. (2011). The neural basis of the right visual field advantage in reading: An MEG analysis using virtual electrodes. *Brain and Language*, *118*(3), 53–71. https://doi.org/10.1016/j.bandl.2010.09.003

Barry, J. G., Yasin, I., & Bishop, D. V. M. (2007). Heritable risk factors associated with language impairments. *Genes, Brain and Behavior*, *6*(1), 66–76. https://doi.org/10.1111/j.1601-183X.2006.00232.x

Bartha-Doering, L., Kollndorfer, K., Kasprian, G., Novak, A., Schuler, A.-L., Fischmeister, F. P. S., … Berl, M. M. (2018). Weaker semantic language lateralization associated with better semantic language performance in healthy right-handed children. *Brain and Behavior*, *8*(11), e01072. https://doi.org/10.1002/brb3.1072

Basic, S., Hajnsek, S., Poljakovic, Z., Basic, M., Culic, V., & Zadro, I. (2004).

Determination of cortical language dominance using functional transcranial Doppler sonography in left-handers. *Clinical Neurophysiology*, *115*(1), 154–160. https://doi.org/10.1016/S1388-2457(03)00281-5

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *1*(1). https://doi.org/10.18637/jss.v067.i01

Bathelt, J., Gathercole, S. E., Butterfield, S., & Astle, D. E. (2018). Children's academic attainment is linked to the global organization of the white matter connectome. *Developmental Science*, *21*(5). https://doi.org/10.1111/desc.12662

Bathelt, J., Holmes, J., Astle, D. E., Holmes, J., Gathercole, S., Astle, D., … Kievit, R. (2018). Data-Driven Subtyping of Executive Function–Related Behavioral Problems in Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(4), 252–262.e4. https://doi.org/10.1016/J.JAAC.2018.01.014

Belger, A., & Banich, M. T. (1998). Costs and benefits of integrating information between the cerebral hemispheres: a computational perspective. *Neuropsychology*, *12*(3), 380–398. https://doi.org/10.1037/0894-4105.12.3.380

Bergey, B. W., Parrila, R. K., & Deacon, H. (2018). Understanding the academic motivations of students with a history of reading difficulty: An expectancy-value-cost approach. *Learning and Individual Differences*, *67*, 41–52. https://doi.org/10.1016/j.lindif.2018.06.008

Berl, M. M., Zimmaro, L. A., Khan, O. I., Dustin, I., Ritzl, E., Duke, E. S., … Gaillard, W. D. (2014). Characterization of atypical language activation patterns in focal epilepsy. *Annals of Neurology*, *75*(1), 33–42. https://doi.org/10.1002/ana.24015

Bernal, B., & Altman, N. R. (2003). Speech Delay in Children: A Functional MR Imaging Study. *Radiology*, *229*(3), 651–658. https://doi.org/10.1148/radiol.2293021746

Bernal, B., & Ardila, A. (2014). Bilateral representation of language: A critical review and analysis of some unusual cases. *Journal of Neurolinguistics*, *28*, 63–80. https://doi.org/10.1016/j.jneuroling.2013.10.002

Bethmann, A., Tempelmann, C., Bleser, R. De, Scheich, H., & Brechmann, A. (2007). Determining language laterality by fMRI and dichotic listening. *Brain Research*, *1133*(1), 145–157. https://doi.org/10.1016/j.brainres.2006.11.057

Binder, J. R., Swanson, S. J., Hammeke, T. A., & Sabsevitz, D. S. (2008). A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia*, *49*(12), 1980–1997. https://doi.org/10.1111/j.1528-1167.2008.01683.x

Bishop, D. V. M. (1990). Handedness and Developmental Disorder. In *Clinics in Developmental Medicine no. 110*. Philadelphia: Blackwell scientific.

Bishop, D. V. M. (2004). *Expression, Reception and Recall of Narrative Instrument (ERRNI)*. London: Pearson Assessment.

Bishop, D. V. M. (2013). Cerebral Asymmetry and Language Development: Cause, Correlate, or Consequence? *Science*, *340*(6138), 1230531. https://doi.org/10.1126/science.1230531

Bishop, D. V. M., Holt, G., Whitehouse, A. J. O., & Groen, M. (2014). No population bias to left-hemisphere language in 4-year-olds with language impairment. *PeerJ*, *2*, e507. https://doi.org/10.7717/peerj.507

Bishop, D. V. M., & Robson, J. (1989). Unimpaired short-term memory and rhyme judgement in congenitally speechless individuals: Implications for the notion of "articulatory coding." *The Quarterly Journal of Experimental Psychology Section A*, *41*(1), 123–140. https://doi.org/http://dx.doi.org/10.1080/14640748908402356

Bishop, D. V. M., Ross, V. A., Daniels, M. S., & Bright, P. (1996). The measurement of hand preference: A validation study comparing three groups of right-handers. *British Journal of Psychology*, *87*(2), 269–285. https://doi.org/10.1111/j.2044-8295.1996.tb02590.x

Bishop, D. V. M., Whitehouse, A., & Sharp, M. (2009). *Communication Checklist - Self Report (CC-SR)*. London: Pearson Assessment.

Boles, D. B., Barth, J. M., & Merrill, E. C. (2008). Asymmetry and performance: Toward a neurodevelopmental theory. *Brain and Cognition*, *66*(2), 124–139. https://doi.org/10.1016/j.bandc.2007.06.002

Bradshaw, A. R., Bishop, D. V. M., & Woodhead, Z. V. J. (2017). Methodological considerations in assessment of language lateralisation with fMRI: A systematic review. *PeerJ*, *2017*(7). https://doi.org/10.7717/peerj.3557

Bradshaw, A. R., Thompson, P. A., Wilson, A. C., Bishop, D. V. M., & Woodhead, Z. V. J. (2017). Measuring language lateralisation with different language tasks: A systematic review. *PeerJ*, *2017*(10). https://doi.org/10.7717/peerj.3929

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/https://doi.org/10.1163/156856897X00357

Branco, D. M., Suarez, R. O., Whalen, S., O'Shea, J. P., Nelson, A. P., da Costa, J. C., & Golby, A. J. (2006). Functional MRI of memory in the hippocampus: Laterality indices may be more meaningful if calculated from whole voxel distributions. *NeuroImage*, *32*(2), 592–602. https://doi.org/10.1016/j.neuroimage.2006.04.201

Brederoo, S. G., Nieuwenstein, M. R., Cornelissen, F. W., & Lorist, M. M. (2019). Reproducibility of visual-field asymmetries: Nine replication studies investigating lateralization of visual information processing. *Cortex*, *111*, 100–126. https://doi.org/10.1016/j.cortex.2018.10.021

Brennan, N. M. P., Whalen, S., de Morales Branco, D., O'Shea, J. P., Norton, I. H., & Golby, A. J. (2007). Object naming is a more sensitive measure of

speech localization than number counting: Converging evidence from direct cortical stimulation and fMRI. *NeuroImage*, *37*(Suppl.1), S100–S108. https://doi.org/10.1016/j.neuroimage.2007.04.052

Broca, P. P. (1861). *Perte de la parole, ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau*. *Bulletin de la Société Anthropologique de Paris* (Vol. 2).

Brown, W. S., Larson, E. B., & Jeeves, M. A. (1994). Directional asymmetries in interhemispheric transmission time: Evidence from visual evoked potentials. *Neuropsychologia*, *32*(4), 439–448. https://doi.org/10.1016/0028-3932(94)90089-2

Bryan, K. L. (1988). Assessment of language disorders after right hemisphere damage. *International Journal of Language & Communication Disorders*, *23*(2), 111–125. https://doi.org/10.3109/13682828809019881

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, *14*(5), 365. https://doi.org/http://dx.doi.org/10.1038/nrn3475

Cabeza, R. (2002). Hemispheric asymmetry reduction in older adults: The HAROLD model. *Psychology and Aging*, *17*(1), 85–100. https://doi.org/10.1037//0882-7974.17.1.85

Carreiras, M., Seghier, M. L., Baquero, S., Estévez, A., Lozano, A., Devlin, J. T., & Price, C. J. (2009). An anatomical signature for literacy. *Nature*, *461*(7266), 983–986. https://doi.org/10.1038/nature08461

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, *47*(2), 149–180. https://doi.org/10.1016/0010-0277(93)90003-E

Cattell, R. B., & Cattell, A. K. S. (1973). *Handbook for the Culture Fair Intelligence Test, Scale 2*. Champaign, Illinois: Institute for Personality and Ability Testing.

Cavalli, E., Casalis, S., El Ahmadi, A., Zira, M., Poracchia-George, F., & Cole, P. (2016). Vocabulary skills are well developed in university students with dyslexia: Evidence from multiple case studies. *Research in Developmental Disabilities*, *51–52*, 89–102. https://doi.org/10.1016/j.ridd.2016.01.006

Cavalli, E., Duncan, L. G., Elbro, C., El Ahmadi, A., & Cole, P. (2017). Phonemic-Morphemic dissociation in university students with dyslexia: an index of reading compensation? *Annals of Dyslexia*, *67*(1), 63–84. https://doi.org/10.1007/s11881-016-0138-y

Chatzitheochari, S., & Platt, L. (2019). Disability differentials in educational attainment in England: primary and secondary effects. *The British Journal of Sociology*, *70*(2), 502–525. https://doi.org/10.1111/1468-4446.12372

Chi, J., Dooling, E., & Gilles, F. (1977). Left-right asymmetries of the temporal speech areas of the human fetus. *Archives of Neurology*, *34*(6), 346–348. https://doi.org/10.1001/archneur.1977.00500180040008

Chiarello, C., & Maxfield, L. (1996). Varieties of interhemispheric inhibition, or

how to keep a good hemisphere down. *Brain and Cognition*, *30*(1), 81–108. https://doi.org/10.1006/brcg.1996.0006

Chiarello, C., Welcome, S. E., Halderman, L. K., & Leonard, C. M. (2009). Does degree of asymmetry relate to performance? An investigation of word recognition and reading in consistent and mixed handers. *Brain and Cognition*, *69*(3), 521–530. https://doi.org/10.1016/J.BANDC.2008.11.002

Clements, A. M., Rimrodt, S. L., Abel, J. R., Blankner, J. G., Mostofsky, S. H., Pekar, J. J., … Cutting, L. E. (2006). Sex differences in cerebral laterality of language and visuospatial processing. *Brain and Language*, *98*(2), 150–158. https://doi.org/https://doi.org/10.1016/j.bandl.2006.04.007

Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the Visual Word Form Area. *Brain*, *125*(5), 1054–1069. https://doi.org/https://doi.org/10.1093/brain/awf094

Corballis, M. C. (2017). The Evolution of Lateralized Brain Circuits. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01021

Corballis, M. C. (2018). Mirror-Image Equivalence and Interhemispheric Mirror-Image Reversal. *Frontiers in Human Neuroscience*, *12*, 140. https://doi.org/10.3389/fnhum.2018.00140

Corballis, M. C., Birse, K., Paggi, A., Manzoni, T., Pierpaoli, C., & Fabri, M. (2010). Mirror-image discrimination and reversal in the disconnected hemispheres. *Neuropsychologia*, *48*(6), 1664–1669. https://doi.org/10.1016/j.neuropsychologia.2010.02.011

Cousin, E., Peyrin, C., Pichat, C., Lamalle, L., Bas, J.-F. Le, & Baciu, M. (2007). Functional MRI approach for assessing hemispheric predominance of regions activated by a phonological and a semantic task. *European Journal of Radiology*, *63*(2), 274–285. https://doi.org/10.1016/j.ejrad.2007.01.030

Daini, R., De Fabritiis, P., Ginocchio, C., Lenti, C., Lentini, C. M., Marzorati, D., & Lorusso, M. L. (2018). Revisiting Strephosymbolie: The Connection between Interhemispheric Transfer and Developmental Dyslexia. *Brain Sciences*, *8*(4). https://doi.org/10.3390/brainsci8040067

Dax, M. (1865). Lésions de la moitié gauche de l'encéphale coïncidant avec l'oubli dessignes de la pensée. *Gazette Hebdomadaire de Médecine et de Chirurgie*, *2*(2), 259–262.

de Guibert, C., Maumet, C., Jannin, P., Ferré, J.-C., Tréguier, C., Barillot, C., … Biraben, A. (2011). Abnormal functional lateralization and activity of language brain areas in typical specific language impairment (developmental dysphasia). *Brain*, *134*(10), 3044–3058. https://doi.org/10.1093/brain/awr141

Deblaere, K., Backes, W. H., Hofman, P., Vandemaele, P., Boon, P. A., Vonck, K., … Aldenkamp, A. (2002). Developing a comprehensive presurgical functional MRI protocol for patients with intractable temporal lobe epilepsy: a pilot study. *Neuroradiology*, *44*(8), 667–673. https://doi.org/10.1007/s00234-002-0800-4

Deppe, M., Knecht, S., Henningsen, H., & Ringelstein, E. . (1997). AVERAGE: a Windows® program for automated analysis of event related cerebral blood flow. *Journal of Neuroscience Methods*, *75*(2), 147–154. https://doi.org/10.1016/S0165-0270(97)00067-8

Dewey, D., Kaplan, B. J., Crawford, S. G., & Wilson, B. N. (2002). Developmental coordination disorder: associated problems in attention, learning, and psychosocial adjustment. *Human Movement Science*, *21*(5–6), 905–18. https://doi.org/10.1016/S0167-9457(02)00163-X

Dodoo-Schittko, F., Rosengarth, K., Doenitz, C., & Greenlee, M. W. (2012). Assessing language dominance with functional MRI: The role of control tasks and statistical analysis. *Neuropsychologia*, *50*(11), 2684–2691. https://doi.org/10.1016/j.neuropsychologia.2012.07.032

Dorsaint-Pierre, R., Penhune, V. B., Watkins, K. E., Neelin, P., Lerch, J. P., Bouffard, M., & Zatorre, R. J. (2006). Asymmetries of the planum temporale and Heschl's gyrus: relationship to language lateralization. *Brain*, *129*(5), 1164–1176. Retrieved from http://dx.doi.org/10.1093/brain/awl055

Dougherty, R. F., Ben-Shachar, M., Deutsch, G. K., Hernandez, A., Fox, G. R., & Wandell, B. A. (2007). Temporal-callosal pathway diffusivity predicts phonological skills in children. *Proceedings of the National Academy of Sciences*, *104*(20), 8556–8561. https://doi.org/10.1073/pnas.0608961104

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Ann. Statist.*, *7*(1), 1–26. https://doi.org/10.1214/aos/1176344552

Eglinton, E., & Annett, M. (1994). Handedness and Dyslexia: A Meta-Analysis. *Perceptual and Motor Skills*, *79*(3_suppl), 1611–1616. https://doi.org/10.2466/pms.1994.79.3f.1611

Ellis, A. W., Ferreira, R., Cathles-Hagan, P., Holt, K., Jarvis, L., & Barca, L. (2009). Word learning and the cerebral hemispheres: from serial to parallel processing of written words. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3675–3696. https://doi.org/https://doi.org/10.1098/rstb.2009.0187

Everts, R., Lidzba, K., Wilke, M., Kiefer, C., Mordasini, M., Schroth, G., … Steinlin, M. (2009). Strengthening of laterality of verbal and visuospatial functions during childhood and adolescence. *Human Brain Mapping*, *30*(2), 473–483. https://doi.org/https://doi.org/10.1002/hbm.20523

Fedorenko, E., Nieto-Castanon, A., & Kanwisher, N. (2012). Lexical and syntactic representations in the brain: An fMRI investigation with multi-voxel pattern analyses. *Neuropsychologia*, *50*(4), 499–513. https://doi.org/10.1016/j.neuropsychologia.2011.09.014

Fernandez, G., de Greiff, A., von Oertzen, J., Reuber, M., Lun, S., Klaver, P., … Elger, C. E. (2001). Language mapping in less than 15 minutes: Real-time functional MRI during routine clinical investigation. *NeuroImage*, *14*(3), 585–594. https://doi.org/10.1006/nimg.2001.0854

Fesl, G., Bruhns, P., Rau, S., Wiesmann, M., Ilmberger, J., Kegel, G., & Brueckmann, H. (2010). Sensitivity and reliability of language laterality

assessment with a free reversed association task-a fMRI study. *European Radiology*, *20*(3), 683–695. https://doi.org/10.1007/s00330-009-1602-4

Finn, E. S., Shen, X., Holahan, J. M., Scheinost, D., Lacadie, C., Papademetris, X., … Constable, R. T. (2014). Disruption of Functional Networks in Dyslexia: A Whole-Brain, Data-Driven Analysis of Connectivity. *Biological Psychiatry*, *76*(5), 397–404. https://doi.org/10.1016/J.BIOPSYCH.2013.08.031

Friederici, A. D. (2011). The Brain Basis of Language Processing: from Structure to Function. *Physiological Reviews*, *91*(4), 1357–1392. https://doi.org/10.1152/physrev.00006.2011

Friedrich, P., Ocklenburg, S., Heins, N., Schlüter, C., Fraenz, C., Beste, C., … Genç, E. (2017). Callosal microstructure affects the timing of electrophysiological left-right differences. *NeuroImage*, *163*, 310–318. https://doi.org/10.1016/J.NEUROIMAGE.2017.09.048

Frye, R. E., Hasan, K., Xue, L., Strickland, D., Malmberg, B., Liederman, J., & Papanicolaou, A. (2008). Splenium microstructure is related to two dimensions of reading skill. *NeuroReport*, *19*(16), 1627–1631. https://doi.org/10.1097/WNR.0b013e328314b8ee

Frässle, S., Krach, S., Paulus, F. M., & Jansen, A. (2016). Handedness is related to neural mechanisms underlying hemispheric lateralization of face processing. *Scientific Reports*, *6*, 27153. https://doi.org/10.1038/srep27153

Gaillard, W. D., Balsamo, L., Xu, B., Grandin, C. B., Braniecki, S. H., Papero, P. H., … Theodore, W. H. (2002). Language dominance in partial epilepsy patients identified with an fMRI reading task. *Neurology*, *59*(2), 256–265. https://doi.org/10.1212/WNL.59.2.256

Gaillard, W. D., Sachs, B. C., Whitnah, J. R., Ahmad, Z., Balsamo, L. M., Petrella, J. R., … Grandin, C. B. (2003). Developmental aspects of language processing: fMRI of verbal fluency in children and adults. *Human Brain Mapping*, *18*(3), 176–185. https://doi.org/10.1002/hbm.10091

Gazzaniga, M. S., & Hillyard, S. (1971). Language and speech capacity of the right hemisphere. *Neuropsychologia*, *9*(3), 273–280. https://doi.org/10.1016/0028-3932(71)90022-4

Gazzaniga, M. S., & Sperry, R. W. (1967). Language after section of the cerebral commissures. *Brain*, *90*(1), 131–148.

Geschwind, N., & Levitsky, W. (1968). Human Brain: Left-Right Asymmetries in Temporal Speech Region. *Science*, *161*(3837), 186–187. https://doi.org/10.1126/science.161.3837.186

Gorgolewski, K. J., Storkey, A. J., Bastin, M. E., & Pernet, C. R. (2012). Adaptive thresholding for reliable topological inference in single subject fMRI analysis. *Frontiers in Human Neuroscience*, *6*, 245. https://doi.org/10.3389/fnhum.2012.00245

Gotts, S. J., Jo, H. J., Wallace, G. L., Saad, Z. S., Cox, R. W., & Martin, A. (2013). Two distinct forms of functional lateralization in the human brain.

Proceedings of the National Academy of Sciences, *110*(36), E3435–E3444. https://doi.org/10.1073/pnas.1302581110

Groen, M. A., Whitehouse, A. J. O., Badcock, N. A., & Bishop, D. V. M. (2012). Does cerebral lateralization develop? A study using functional transcranial Doppler ultrasound assessing lateralization for language production and visuospatial memory. *Brain and Behavior*, *2*(3), 256–269. https://doi.org/10.1002/brb3.56

Groen, M. A., Whitehouse, A. J. O., Badcock, N. A., & Bishop, D. V. M. (2013). Associations between Handedness and Cerebral Lateralisation for Language: A Comparison of Three Measures in Children. *PloS ONE*, *8*(5), 1–6. https://doi.org/10.1371/journal.pone.0064876

Häberling, I. S., Badzakova-Trajkov, G., & Corballis, M. C. (2011). Callosal tracts and patterns of hemispheric dominance: A combined fMRI and DTI study. *NeuroImage*, *54*(2), 779–786. https://doi.org/10.1016/J.NEUROIMAGE.2010.09.072

Häberling, I. S., Steinemann, A., & Corballis, M. C. (2016). Cerebral asymmetry for language: Comparing production with comprehension. *Neuropsychologia*, *80*, 17–23. https://doi.org/10.1016/j.neuropsychologia.2015.11.002

Hakvoort, B., van der Leij, A., van Setten, E., Maurits, N., Maassen, B., & van Zuijen, T. (2016). Dichotic listening as an index of lateralization of speech perception in familial risk children with and without dyslexia. *Brain and Cognition*, *109*, 75–83. https://doi.org/10.1016/J.BANDC.2016.09.004

Hammill, D. D., Brown, V. L., Larson, S. C., & Wiederholt, J. L. (2007). *Test of adolescent and adult language*. Austin: Pro-Ed.

Harrington, G. S., Buonocore, M. H., & Farias, S. T. (2006). Intrasubject reproducibility of functional MR imaging activation in language tasks. *American Journal of Neuroradiology*, *27*(4), 938–944.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hastie, T., Tibshirani, R., & Friedman, J. (2001). Unsupervised learning. In *The elements of statistical learning* (pp. 485–586). New York: Springer New York Inc. https://doi.org/10.1007/978-0-387-84858-7

Hatcher, J., Snowling, M. J., & Griffiths, Y. M. (2002). Cognitive assessment of dyslexic students in higher education. *British Journal of Educational Psychology*, *72*(1), 119–133. https://doi.org/10.1348/000709902158801

Hebb, D. (1949). *The organization of behavior. A neuropsychological theory*. New York: Wiley.

Hellige, J. B. (1993). *Hemispheric asymmetry: what's right and what's left*. *Perspectives in Cognitive Neuroscience*. Cambridge, MA: Harvard

University Press.

Henderson, L., Barca, L., & Ellis, A. W. (2007). Interhemispheric cooperation and non-cooperation during word recognition: Evidence for callosal transfer dysfunction in dyslexic adults. *Brain and Language*, *103*(3), 276–291. https://doi.org/https://doi.org/10.1016/j.bandl.2007.04.009

Henderson, P. (2017). Are there delays in reporting dyslexia in university learners? Experiences of university learning support staff. *Journal of Further and Higher Education*, *41*(1), 30–43. https://doi.org/10.1080/0309877X.2015.1023563

Hernandez, N., Andersson, F., Edjlali, M., Hommet, C., Cottier, J. P., Destrieux, C., & Bonnet-Brilhault, F. (2013). Cerebral functional asymmetry and phonological performance in dyslexic adults. *Psychophysiology*, *50*(12), 1226–1238. https://doi.org/10.1111/psyp.12141

Heron, W. (1957). Perception as a function of retinal locus and attention. *American Journal of Psychology*, *70*(1), 38–48. https://doi.org/10.2307/1419227

Hickok, G., & Poeppel, D. (2007). Opinion - The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. https://doi.org/10.1038/nrn2113

Hillis, A. E. (2007). Aphasia - Progress in the last quarter of a century. *Neurology*, *69*(2), 200–213. https://doi.org/10.1212/01.wnl.0000265600.69385.6f

Hirnstein, M., Hugdahl, K., & Hausmann, M. (2014). How brain asymmetry relates to performance - A large-scale dichotic listening study. *Frontiers in Psychology*, *5*, 58. https://doi.org/10.3389/fpsyg.2014.00058

Hirnstein, M., Leask, S., Rose, J., & Hausmann, M. (2010). Disentangling the relationship between hemispheric asymmetry and cognitive performance. *Brain and Cognition*, *73*(2), 119–127. https://doi.org/10.1016/J.BANDC.2010.04.002

Holland, S., Plante, E., Byars, A., H. Strawsburg, R., J. Schmithorst, V., & Ball, W. (2001). Normal fMRI Brain Activation Patterns in Children Performing a Verb Generation Task. *NeuroImage*, *14*(4), 837–843. https://doi.org/10.1006/nimg.2001.0875

Hong, D. S. (2018). “Communities” of Conditions: Novel Methods for Classifying Psychiatric Disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *57*(4), 233–234. https://doi.org/10.1016/j.jaac.2018.02.002

Hope, T. M. H., Prejawa, S., Parker Jones, ‘Ōiwi, Oberhuber, M., Seghier, M. L., Green, D. W., & Price, C. J. (2014). Dissecting the functional anatomy of auditory word repetition. *Frontiers in Human Neuroscience*, *8*, 246. https://doi.org/10.3389/fnhum.2014.00246

Horowitz, A., Barazany, D., Tavor, I., Bernstein, M., Yovel, G., & Assaf, Y. (2015). In vivo correlation between axon diameter and conduction velocity

in the human brain. *Brain Structure and Function*, *220*(3), 1777–1788. https://doi.org/10.1007/s00429-014-0871-0

Horwitz, B., Rumsey, J. M., & Donohue, B. C. (1998). Functional connectivity of the angular gyrus in normal reading and dyslexia. *Proceedings of the National Academy of Sciences*, *95*(15), 8939–8944. https://doi.org/10.1073/pnas.95.15.8939

Hugdahl, K. (2011). Fifty years of dichotic listening research – Still going and going and…. *Brain and Cognition*, *76*(2), 211–213. https://doi.org/10.1016/J.BANDC.2011.03.006

Hund-Georgiadis, M., Lex, U., Friederici, A. D., & von Cramon, D. Y. (2002). Non-invasive regime for language lateralization in right- and left-handers by means of functional MRI and dichotic listening. *Experimental Brain Research*, *145*(2), 166–176. https://doi.org/10.1007/s00221-002-1090-0

Hund-Georgiadis, M., Lex, U., & von Cramon, D. Y. (2001). Language dominance assessment by means of fMRI: Contributions from task design, performance, and stimulus modality. *Journal of Magnetic Resonance Imaging*, *13*(5), 668–675. https://doi.org/10.1002/jmri.1094

Hunter, Z. R., & Brysbaert, M. (2008). Visual half-field experiments are a good measure of cerebral language dominance if used properly: Evidence from fMRI. *Neuropsychologia*, *46*(1), 316–325. https://doi.org/10.1016/j.neuropsychologia.2007.07.007

Hynd, G., Hall, J., Novey, E., & et al. (1995). Dyslexia and corpus callosum morphology. *Archives of Neurology*, *52*(1), 32–38. https://doi.org/10.1001/archneur.1995.00540250036010

Ide, A., Dolezal, C., Fernández, M., Labbé, E., Mandujano, R., Montes, S., … Aboitiz, F. (1999). Hemispheric differences in variability of fissural patterns in parasylvian and cingulate regions of human brains. *Journal of Comparative Neurology*, *410*(2), 235–242. https://doi.org/10.1002/(SICI)1096-9861(19990726)410:2<235::AID-CNE5>3.0.CO;2-G

Illingworth, S., & Bishop, D. V. M. (2009). Atypical cerebral lateralisation in adults with compensated developmental dyslexia demonstrated using functional transcranial Doppler ultrasound. *Brain and Language*, *111*(1), 61–65. https://doi.org/10.1016/j.bandl.2009.05.002

Iwabuchi, S. J., & Kirk, I. J. (2009). Atypical interhemispheric communication in left-handed individuals. *NeuroReport*, *20*(2), 166–169. https://doi.org/10.1097/WNR.0b013e32831f1cbb

Jansen, A., Menke, R., Sommer, J., Förster, A. F., Bruchmann, S., Hempleman, J., … Knecht, S. (2006). The assessment of hemispheric lateralization in functional MRI—Robustness and reproducibility. *NeuroImage*, *33*(1), 204–217. https://doi.org/10.1016/j.neuroimage.2006.06.019

Jensen-Kondering, U. R., Ghobadi, Z., Wolff, S., Jansen, O., & Ulmer, S. (2012). Acoustically presented semantic decision-making tasks provide a robust depiction of the temporo-parietal speech areas. *Journal of Clinical*

Neuroscience, *19*(3), 428–433. https://doi.org/10.1016/j.jocn.2011.04.038

Joliot, M., Tzourio-Mazoyer, N., & Mazoyer, B. (2016). Intra-hemispheric intrinsic connectivity asymmetry and its relationships with handedness and language Lateralization. *Neuropsychologia*, *93*, 437–447. https://doi.org/10.1016/j.neuropsychologia.2016.03.013

Josse, G., Seghier, M. L., Kherif, F., & Price, C. J. (2008). Explaining Function with Anatomy: Language Lateralization and Corpus Callosum Size. *Journal of Neuroscience*, *28*(52), 14132–14139. https://doi.org/10.1523/JNEUROSCI.4383-08.2008

Kadesjö, B., & Gillberg, C. (1999). Developmental coordination disorder in Swedish 7-year-old children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*(7), 820–828. https://doi.org/10.1097/00004583-199907000-00011

Karolis, V. R., Corbetta, M., & Thiebaut de Schotten, M. (2019). The architecture of functional lateralisation and its relationship to callosal connectivity in the human brain. *Nature Communications*, *10*(1), 1417. https://doi.org/10.1038/s41467-019-09344-1

Keller, C., & Kell, C. A. (2016). Asymmetric intra- and interhemispheric interactions during covert and overt sentence reading. *Neuropsychologia*, *93*(SI), 448–465. https://doi.org/10.1016/j.neuropsychologia.2016.04.002

Kelly, M. L., Jones, M. W., McDonald, S. A., & Shillcock, R. C. (2004). Dyslexics' eye fixations may accommodate to hemispheric desynchronization. *NeuroReport*, *15*(17), 2629–2632. https://doi.org/10.1097/00001756-200412030-00014

Kennan, R. P., Kim, D., Maki, A., Koizumi, H., & Constable, R. T. (2002). Non-invasive assessment of language lateralization by Transcranial near infrared optical topography and functional MRI. *Human Brain Mapping*, *16*(3), 183–189. https://doi.org/10.1002/hbm.10039

Kherdr, E. M., Hamed, E., Said, A., & Basahi, J. (2002). Handedness and language cerebral lateralization. *European Journal of Applied Physiology*, *87*(4–5), 469–473. https://doi.org/10.1007/s00421-002-0652-y

Kim, J., Kim, Y. W., Park, C., Park, E. S., Kim, H.-H., Lee, S.-K., & Kim, D. I. (2006). Diffusion-tensor magnetic resonance imaging in children with language impairment. *NeuroReport*, *17*(12), 1279–1282. https://doi.org/10.1097/01.wnr.0000230516.86090.67

Kimura, D. (1961). Cerebral Dominance and the Perception of Verbal Stimuli. *Canadian Journal of Psychology*, *15*(3), 166–171. https://doi.org/10.1037/h0083219

Kinsbourne, M. (1970). The cerebral basis of lateral asymmetries in attention. *Acta Psychologica*, *33*, 193–201. https://doi.org/10.1016/0001-6918(70)90132-0

Kirby, A., Sugden, D., Beveridge, S., Edwards, L., & Edwards, R. (2008). Dyslexia and developmental co-ordination disorder in further and higher

education - Similarities and differences. Does the `label' influence the support given? *Dyslexia*, *14*(3), 197–213. https://doi.org/10.1002/dys.367

Kleinhans, N. M., Mueller, R.-A., Cohen, D. N., & Courchesne, E. (2008). Atypical functional lateralization of language in autism spectrum disorders. *Brain Research*, *1221*, 115–125. https://doi.org/10.1016/j.brainres.2008.04.080

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., … Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, *123*(12), 2512–2518. https://doi.org/https://doi.org/10.1093/brain/123.12.2512

Knecht, S., Dräger, B., Flöel, A., Lohmann, H., Breitenstein, C., Deppe, M., … Ringelstein, E.-B. (2001). Behavioural relevance of atypical language lateralization in healthy subjects. *Brain*, *124*(8), 1657–1665. https://doi.org/10.1093/brain/124.8.1657

Knecht, S., Henningsen, H., Deppe, M., Huber, T., Ebner, A., & Ringelstein, E. B. (1996). Successive activation of both cerebral hemispheres during cued word generation. *Neuroreport*, *7*(3), 820–824. https://doi.org/10.1097/00001756-199602290-00033

Knecht, S., Jansen, A., Frank, A., van Randenborgh, J., Sommer, J., Kanowski, M., & Heinze, H. . (2003). How atypical is atypical language dominance? *NeuroImage*, *18*(4), 917–927. https://doi.org/10.1016/S1053-8119(03)00039-9

Kohler, M., Keage, H. A. D., Spooner, R., Flitton, A., Hofmann, J., Churches, O. F., … Badcock, N. A. (2015). Variability in lateralised blood flow response to language is associated with language development in children aged 1-5 years. *Brain and Language*, *145*, 34–41. https://doi.org/10.1016/j.bandl.2015.04.004

Korkman, M., Kirk, U., & Kemp, S. I. (1998). *NEPSY: a developmental neuropsychological assessment*. San Antonio: Psychological Corporation.

Kosslyn, S. M. (1987). Seeing and Imagining in the Cerebral Hemispheres: A Computational Approach. *Psychological Review*, *94*(2), 148–175. https://doi.org/10.1037/0033-295X.94.2.148

Krach, S., Chen, L. M., & Hartje, W. (2006). Comparison between visual half-field performance and cerebral blood flow changes as indicators of language dominance. *Laterality: Asymmetries of Body, Brain and Cognition*, *11*(2), 122–140. https://doi.org/https://doi.org/10.1080/13576500500384975

Lee, D., Swanson, S. J., Sabsevitz, D. S., Hammeke, T. A., Winstanley, F. S., Possing, E. T., & Binder, J. R. (2008). Functional MRI and Wada studies in patients with interhemispheric dissociation of language functions. *Epilepsy & Behavior*, *13*(2), 350–356. https://doi.org/10.1016/j.yebeh.2008.04.010

Lefly, D. L., & Pennington, B. F. (1991). Spelling-errors and reading fluency in compensated adult dyslexics. *Annals of Dyslexia*, *41*(1), 143–162. https://doi.org/10.1007/BF02648083

Leroy, F., Cai, Q., Bogart, S. L., Dubois, J., Coulon, O., Monzalvo, K., … Dehaene-Lambertz, G. (2015). New human-specific brain landmark: The depth asymmetry of superior temporal sulcus. *Proceedings of the National Academy of Sciences*, *112*(4), 1208–1213. https://doi.org/10.1073/pnas.1412389112

Li, B., & Freeman, R. D. (2015). Neurometabolic coupling between neural activity, glucose, and lactate in activated visual cortex. *Journal of Neurochemistry*, *135*(4), 742–754. https://doi.org/10.1111/jnc.13143

Liberman, A. M., & Mattingly, I. G. (1985). The Motor Theory of Speech-Perception Revised. *Cognition*, *21*(1), 1–36. https://doi.org/10.1016/0010-0277(85)90021-6

Lindell, A. K., & Hudry, K. (2013). Atypicalities in Cortical Structure, Handedness, and Functional Lateralization for Language in Autism Spectrum Disorders. *Neuropsychology Review*, *23*(3), 257–270. https://doi.org/10.1007/s11065-013-9234-5

Liu, H., Stufflebeam, S. M., Sepulcre, J., Hedden, T., & Buckner, R. L. (2009). Evidence from intrinsic activity that asymmetry of the human brain is controlled by multiple factors. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(48), 20499–503. https://doi.org/10.1073/pnas.0908073106

Markus, H., Pereira, A., & Cloud, G. (2016). Vascular anatomy and stroke syndromes. In *Stroke Medicine (Oxford Specialist Handbooks in Neurology)* (2nd ed.). Oxford, UK. https://doi.org/10.1093/med/9780198737889.003.0003

Martin, C. D., Thierry, G., Démonet, J. F., Roberts, M., & Nazir, T. (2007). ERP evidence for the split fovea theory. *Brain Research*, *1185*, 212–220. https://doi.org/10.1016/j.brainres.2007.09.049

Marzi, C. A., Bisiacchi, P., & Nicoletti, R. (1991). Is interhemispheric transfer of visuomotor information asymmetric? Evidence from a meta-analysis. *Neuropsychologia*, *29*(12), 1163–1177. https://doi.org/10.1016/0028-3932(91)90031-3

Mazoyer, B., Mellet, E., Perchey, G., Zago, L., Crivello, F., Jobard, G., … TzourioMazoyer, N. (2016). BIL&GIN: A neuroimaging, cognitive, behavioral, and genetic database for the study of human brain lateralization. *NeuroImage*, *124*(Part B), 1225–1231. https://doi.org/https://doi.org/10.1016/j.neuroimage.2015.02.071

Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Perchey, G., … Tzourio-Mazoyer, N. (2014). Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PloS ONE*, *9*(6). https://doi.org/10.1371/journal.pone.0101165

Mellet, E., Zago, L., Jobard, G., Crivello, F., Petit, L., Joliot, M., … Tzourio-Mazoyer, N. (2014). Weak language lateralization affects both verbal and spatial skills: An fMRI study in 297 subjects. *Neuropsychologia*, *65*, 56–62.

https://doi.org/10.1016/j.neuropsychologia.2014.10.010

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247–279. https://doi.org/10.1016/0010-0285(82)90010-X

Minagawa-Kawai, Y., Cristià, A., & Dupoux, E. (2011). Cerebral lateralization and early speech acquisition: A developmental scenario. *Developmental Cognitive Neuroscience*, *1*(3), 217–232. https://doi.org/10.1016/j.dcn.2011.03.005

Minagawa-Kawai, Y., Mori, K., Naoi, N., & Kojima, S. (2007). Neural Attunement Processes in Infants during the Acquisition of a Language-Specific Phonemic Contrast. *Journal of Neuroscience*, *27*(2), 315–321. https://doi.org/10.1523/JNEUROSCI.1984-06.2007

Miro, J., Ripolles, P., Lopez-Barroso, D., Vila-Ballo, A., Juncadella, M., de Diego-Balaguer, R., … Falip, M. (2014). Atypical language organization in temporal lobe epilepsy revealed by a passive semantic paradigm. *Bmc Neurology*, *14*, 98. https://doi.org/10.1186/1471-2377-14-98

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PloS Med*, *6*(7), e1000097. https://doi.org/https://doi.org/10.1371/journal.pmed.1000097

Mohr, B., Endrass, T., Hauk, O., & Pulvermüller, F. (2007). ERP correlates of the bilateral redundancy gain for words. *Neuropsychologia*, *45*(9), 2114–2124. https://doi.org/10.1016/j.neuropsychologia.2007.01.015

Mohr, B., & Pulvermüller, F. (2002). Redundancy Gains and Costs in Cognitive Processing: Effects of Short Stimulus Onset Asynchronies. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(6), 1200–1223. https://doi.org/10.1037/0278-7393.28.6.1200

Mohr, B., Pulvermüller, F., Cohen, R., & Rockstroh, B. (2000). Interhemispheric cooperation during word processing: evidence for callosal transfer dysfunction in schizophrenic patients. *Schizophrenia Research*, *46*(2–3), 231–239. https://doi.org/10.1016/S0920-9964(00)00020-7

Mohr, B., Pulvermüller, F., Rayman, J., & Zaidel, E. (1994). Interhemispheric cooperation during lexical processing is mediated by the corpus callosum: Evidence from the split-brain. *Neuroscience Letters*, *181*(1–2), 17–21. https://doi.org/10.1016/0304-3940(94)90550-9

Monaghan, P., & Shillcock, R. (2008). Hemispheric dissociation and dyslexia in a computational model of reading. *Brain and Language*, *107*(3), 185–193. https://doi.org/10.1016/J.BANDL.2007.12.005

Moore, L. H., Brown, W. S., Markee, T. E., Theberge, D. C., & Zvi, J. C. (1995). Bimanual coordination in dyslexic adults. *Neuropsychologia*, *33*(6), 781–793. https://doi.org/10.1016/0028-3932(95)00019-Y

Moreau, D., Stonyer, J. E., McKay, N. S., & Waldie, K. E. (2018). No evidence for systematic white matter correlates of dyslexia: An Activation Likelihood

Estimation meta-analysis. *Brain Research*, *1683*, 36–47. https://doi.org/10.1016/j.brainres.2018.01.014

Morrison, M. A., Churchill, N. W., Cusimano, M. D., Schweizer, T. A., Das, S., & Graham, S. J. (2016). Reliability of task-based fMRI for preoperative planning: A test-retest study in brain tumor patients and healthy controls. *PLoS ONE*, *11*(2). https://doi.org/10.1371/journal.pone.0149547

Mortimore, T., & Crozier, W. R. (2006). Dyslexia and difficulties with study skills in higher education. *Studies in Higher Education*, *31*(2), 235–251. https://doi.org/10.1080/03075070600572173

Musiek, F. E., & Weihing, J. (2011). Perspectives on dichotic listening and the corpus callosum. *Brain and Cognition*, *76*(2), 225–232. https://doi.org/10.1016/j.bandc.2011.03.011

Nadkarni, T. N., Andreoli, M. J., Nair, V. A., Yin, P., Young, B. M., Kundu, B., … Prabhakaran, V. (2015). Usage of fMRI for pre-surgical planning in brain tumor and vascular lesion patients: Task and statistical threshold effects on language lateralization. *Neuroimage-Clinical*, *7*, 415–423. https://doi.org/10.1016/j.nicl.2014.12.014

Nagata, S., Uchimura, K., Hirakawa, W., & Kuratsu, J. (2001). Method for quantitatively evaluating the lateralization of linguistic function using functional MR imaging. *American Journal of Neuroradiology*, *22*(5), 985–991.

Niskanen, E., Kononen, M., Villberg, V., Nissi, M., Ranta-aho, P., Saisanen, L., … Vanninen, R. (2012). The effect of fMRI task combinations on determining the hemispheric dominance of language functions. *Neuroradiology*, *54*(4), 393–405. https://doi.org/10.1007/s00234-011-0959-7

Nordhausen, K., & Oja, H. (2015). Multivariate L1 Methods: The Package MNM. *Journal of Statistical Software*, *43*(5). https://doi.org/10.18637/jss.v043.i05

Northam, G. B., Liégeois, F., Tournier, J. D., Croft, L. J., Johns, P. N., Chong, W. K., … Baldeweg, T. (2012). Interhemispheric temporal lobe connectivity predicts language impairment in adolescents born preterm. *Brain*, *135*(12), 3781–3798. https://doi.org/10.1093/brain/aws276

Nowicka, A., Grabowska, A., & Fersten, E. (1996). Interhemispheric transmission of information and functional asymmetry of the human brain. *Neuropsychologia*, *34*(2), 147–151. https://doi.org/10.1016/0028-3932(95)00064-X

Nowicka, A., & Tacikowski, P. (2011). Transcallosal transfer of information and functional asymmetry of the human brain. *Laterality*, *16*(1), 35–74. https://doi.org/10.1080/13576500903154231

Ocklenburg, S., Hugdahl, K., & Westerhausen, R. (2013). Structural white matter asymmetries in relation to functional asymmetries during speech perception and production. *NeuroImage*, *83*, 1088–1097. https://doi.org/https://doi.org/10.1016/j.neuroimage.2013.07.076

Oja, H., & Randles, R. H. (2005). Multivariate Nonparametric Tests. *Statistical Science*, *19*(4), 598–605. https://doi.org/10.1214/088342304000000558

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/https://doi.org/10.1016/0028-3932(71)90067-4

Olivares, R., Montiel, J., & Aboitiz, F. (2001). Species differences and similarities in the fine structure of the mammalian corpus callosum. *Brain, Behavior and Evolution*, *57*(2), 98—105. https://doi.org/10.1159/000047229

Orton, S. (1937). *Reading, writing and speech problems in children.* New York: Norton.

Partovi, S., Jacobi, B., Rapps, N., Zipp, L., Karimi, S., Rengier, F., … Stippich, C. (2012). Clinical Standardized fMRI Reveals Altered Language Lateralization in Patients with Brain Tumor. *American Journal of Neuroradiology*, *33*(11), 2151–2157. https://doi.org/10.3174/ajnr.A3137

Partovi, S., Konrad, F., Karimi, S., Rengier, F., Lyo, J. K., Zipp, L., … Stippich, C. (2012). Effects of Covert and Overt Paradigms in Clinical Language fMRI. *Academic Radiology*, *19*(5), 518–525. https://doi.org/10.1016/j.acra.2011.12.017

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*(12), 976–987. https://doi.org/https://doi.org/10.1038/nrn2277

Paul, L. K. (2011). Developmental malformation of the corpus callosum: a review of typical callosal development and examples of developmental disorders with callosal involvement. *Journal of Neurodevelopmental Disorders*, *3*(1), 3–27. https://doi.org/10.1007/s11689-010-9059-y

Paulesu, E., Frith, U., Snowling, M., Gallagher, A., Morton, J., Frackowiak, R. S. J., & Frith, C. D. (1996). Is developmental dyslexia a disconnection syndrome? Evidence from PET scanning. *Brain*, *119*(1), 143–157. https://doi.org/https://doi.org/10.1093/brain/119.1.143

Pedersen, H. F., Fusaroli, R., Lauridsen, L. L., & Parrila, R. (2016). Reading Processes of University Students with Dyslexia - An Examination of the Relationship between Oral Reading and Reading Comprehension. *Dyslexia*, *22*(4), 305–321. https://doi.org/10.1002/dys.1542

Peelle, J. (2012). The hemispheric lateralization of speech processing depends on what “speech” is: a hierarchical perspective. *Frontiers in Human Neuroscience*, *6*, 309. https://doi.org/10.3389/fnhum.2012.00309

Perlaki, G., Horvath, R., Orsi, G., Aradi, M., Auer, T., Varga, E., … Janszky, J. (2013). White-matter microstructure and language lateralization in left-handers: A whole-brain MRI analysis. *Brain and Cognition*, *82*(3), 319–328. https://doi.org/10.1016/j.bandc.2013.05.005

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as “asymmetric sampling in time.” *Speech*

Communication, *41*(1), 245–255. https://doi.org/10.1016/S0167-6393(02)00107-3

Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology*, *28*, 142–149. https://doi.org/10.1016/j.conb.2014.07.005

Polatajko, H., Fox, M., & Missiuna, C. (1995). An International Consensus on Children with Developmental Coordination Disorder. *Canadian Journal of Occupational Therapy*, *62*(1), 3–6. https://doi.org/10.1177/000841749506200101

Pravata, E., Sestieri, C., Mantini, D., Briganti, C., Colicchio, G., Marra, C., … Caulo, M. (2011). Functional Connectivity MR Imaging of the Language Network in Patients with Drug-Resistant Epilepsy. *American Journal of Neuroradiology*, *32*(3), 532–540. https://doi.org/10.3174/ajnr.A2311

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, *62*(2), 816–847. https://doi.org/10.1016/j.neuroimage.2012.04.062

Price, C. J., & Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246–253. https://doi.org/10.1016/j.tics.2011.04.001

Propper, R. E., O'Donnell, L. J., Whalen, S., Tie, Y., Norton, I. H., Suarez, R. O., … Golby, A. J. (2010). A combined fMRI and DTI examination of functional language lateralization and arcuate fasciculus structure: Effects of degree versus direction of hand preference. *Brain and Cognition*, *73*(2), 85–92. https://doi.org/https://doi.org/10.1016/j.bandc.2010.03.004

Putnam, M. C., Wig, G. S., Grafton, S. T., Kelley, W. M., & Gazzaniga, M. S. (2008). Structural organization of the corpus callosum predicts the extent and impact of cortical activity in the nondominant hemisphere. *Journal of Neuroscience*, *28*(11), 2912–2918. https://doi.org/10.1523/JNEUROSCI.2295-07.2008

R Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org

Raemaekers, M., Schellekens, W., Petridou, N., & Ramsey, N. F. (2018). Knowing left from right: asymmetric functional connectivity during resting state. *Brain Structure and Function*, *223*(4), 1909–1922. https://doi.org/10.1007/s00429-017-1604-y

Ramsey, N. F., Sommer, I. E. C., Rutten, G. J., & Kahn, R. S. (2001). Combined analysis of language tasks in fMRI improves assessment of hemispheric dominance for language functions in individual subjects. *NeuroImage*, *13*(4), 719–733. https://doi.org/10.1006/nimg.2000.0722

Rasmussen, T., & Milner, B. (1975). Clinical and surgical studies of the cerebral speech areas in man. In K. J. Zulch, O. Creutzfeldt, & G. C. Galbraith (Eds.) (p. 238). Berlin: Springer.

Razafimandimby, A., Maiza, O., Herve, P.-Y., Lecardeur, L., Delamillieure, P., Brazo, P., … Dollfus, S. (2007). Stability of functional language lateralization over time in schizophrenia patients. *Schizophrenia Research*, *94*(1–3), 197–206. https://doi.org/10.1016/j.schres.2007.04.011

Rice, G. E., Ralph, M. A. L., & Hoffman, P. (2015). The roles of left versus right anterior temporal lobes in conceptual knowledge: An ALE meta-analysis of 97 functional neuroimaging studies. *Cerebral Cortex*, *25*(11), 4374–4391. https://doi.org/10.1093/cercor/bhv024

Ringo, J. L., Doty, R. W., Demeter, S., & Simard, P. Y. (1994). Time Is of the Essence: A Conjecture that Hemispheric Specialization Arises from Interhemispheric Conduction Delay. *Cerebral Cortex*, *4*(4), 331–343. https://doi.org/10.1093/cercor/4.4.331

Rogers, L. J., Zucca, P., & Vallortigara, G. (2004). Advantages of Having a Lateralized Brain. *Proceedings: Biological Sciences*, *271*, S420–S422. https://doi.org/https://doi.org/10.1098/rsbl.2004.0200

Ruff, I. M., Brennan, N. M. P., Peck, K. K., Hou, B. L., Tabar, V., Brennan, C. W., & Holodny, A. I. (2008). Assessment of the language laterality index in patients with brain tumor using functional MR imaging: Effects of thresholding, task selection, and prior surgery. *American Journal of Neuroradiology*, *29*(3), 528–535. https://doi.org/10.3174/ajnr.A0841

Rutten, G. J. M., Ramsey, N. F., van Rijen, P. C., & van Veelen, C. W. M. (2002). Reproducibility of fMRI-determined language lateralization in individual subjects. *Brain and Language*, *80*(3), 421–437. https://doi.org/10.1006/brln.2001.2600

Sanjuan, A., Bustamante, J.-C., Forn, C., Ventura-Campos, N., Barros-Loscertales, A., Martinez, J.-C., … Avila, C. (2010). Comparison of two fMRI tasks for the evaluation of the expressive language function. *Neuroradiology*, *52*(5), 407–415. https://doi.org/10.1007/s00234-010-0667-8

Sanjuan, A., Forn, C., Ventura-Campos, N., Rodriguez-Pujadas, A., Garcia-Porcar, M., Belloch, V., … Avila, C. (2010a). The sentence verification task: a reliable fMRI protocol for mapping receptive language in individual subjects. *European Radiology*, *20*(10), 2432–2438. https://doi.org/10.1007/s00330-010-1814-7

Sanjuan, A., Forn, C., Ventura-Campos, N., Rodriguez-Pujadas, A., Garcia-Porcar, M., Belloch, V., … Avila, C. (2010b). The sentence verification task: a reliable fMRI protocol for mapping receptive language in individual subjects. *European Radiology*, *20*(10), 2432–2438. https://doi.org/10.1007/s00330-010-1814-7

Sato, Y., Mori, K., Koizumi, T., Minagawa-Kawai, Y., Tanaka, A., Ozawa, E., … Mazuka, R. (2011). Functional Lateralization of Speech Processing in Adults and Children Who Stutter. *Frontiers in Psychology*, *2*, 70. https://doi.org/10.3389/fpsyg.2011.00070

Sattler, J. M., & Dumont, R. (2004). *Assessment of children: WISC-IV and*

WPPSI-III supplement. Jerome M. Sattler, Publisher. San Diego: Jerome M. Satler.

Schulte, T., Sullivan, E. V, Müller-Oehring, E. M., Adalsteinsson, E., & Pfefferbaum, A. (2005). Corpus Callosal Microstructural Integrity Influences Interhemispheric Processing: A Diffusion Tensor Imaging Study. *Cerebral Cortex*, *15*(9), 1384–1392. https://doi.org/10.1093/cercor/bhi020

Schumacke, R. E. (2016). Hotelling's T square: A Two-Group Multivariate Analysis. In *Using R With Multivariate Statistics* (pp. 27–55). Thousand Oaks, CA: Sage Publications.

Seghier, M. L. (2008). Laterality index in functional MRI: methodological issues. *Magnetic Resonance Imaging*, *26*(5), 594–601. https://doi.org/10.1016/j.mri.2007.10.010

Seghier, M. L., Josse, G., Leff, A. P., & Price, C. J. (2011). Lateralization is Predicted by Reduced Coupling from the Left to Right Prefrontal Cortex during Semantic Decisions on Written Words. *Cerebral Cortex*, *21*(7), 1519–1531. https://doi.org/10.1093/cercor/bhq203

Seghier, M. L., Kherif, F., Josse, G., & Price, C. J. (2011). Regional and Hemispheric Determinants of Language Laterality: Implications for Preoperative fMRI. *Human Brain Mapping*, *32*(10), 1602–1614. https://doi.org/10.1002/hbm.21130

Seghier, M. L., Lazeyras, F., Pegna, A. J., Annoni, J.-M., Zimine, I., Mayer, E., … Khateb, A. (2004). Variability of fMRI activation during a phonological and semantic language task in healthy subjects. *Human Brain Mapping*, *23*(3), 140–155. https://doi.org/10.1002/hbm.20053

Seghier, M. L., & Price, C. J. (2013). Dissociating frontal regions that co-lateralize with different ventral occipitotemporal regions during word processing. *Brain and Language*, *126*(2), 133–140. https://doi.org/10.1016/j.bandl.2013.04.003

Sepeta, L. N., Berl, M. M., Wilke, M., You, X., Mehta, M., Xu, B., … Gaillard, W. D. (2016). Age-dependent mesial temporal lobe lateralization in language fMRI. *Epilepsia*, *57*(1), 122–130. https://doi.org/10.1111/epi.13258

Sirkiä, S., Taskinen, S., Nevalainen, J., & Oja, H. (2007). Multivariate nonparametrical methods based on spatial signs and ranks: The R package SpatialNP. Retrieved from https://pdfs.semanticscholar.org/8f07/f275f80e504cf0bd78af992c0be1141c1c4e.pdf

Smythe, P., & Annett, M. (2006). Phonology and handedness in primary school: predictions of the right shift theory. *Journal of Child Psychology and Psychiatry*, *47*(2), 205–212. https://doi.org/10.1111/j.1469-7610.2005.01463.x

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia*, *7*(1), 37–46. https://doi.org/10.1002/dys.185

Snowling, M. J. (2005). The Science of Dyslexia: A Review of Contemporary

Approaches. In *The Study of Dyslexia* (pp. 77–90). https://doi.org/10.1007/0-306-48534-6_4

Sommer, I. E. C., Ramsey, N. F., Mandl, R. C. W., & Kahn, R. S. (2003). Language lateralization in female patients with schizophrenia: an fMRI study. *Schizophrenia Research*, *60*(2–3), 183–190. https://doi.org/10.1016/S0920-9964(02)00300-6

Sperry, R. (1982). Some effects of disconnecting the cerebral hemispheres - Nobel lecture, 8 December 1981. *Bioscience Reports*, *2*(5), 265–276. https://doi.org/10.1007/BF01115112

Steinmetz, H. (1996). Structure, Function and Cerebral Asymmetry: In Vivo Morphometry of the Planum Temporale. *Neuroscience & Biobehavioral Reviews*, *20*(4), 587–591. https://doi.org/10.1016/0149-7634(95)00071-2

Stephan, K. E., Fink, G. R., & Marshall, J. C. (2007). Mechanisms of hemispheric specialization: Insights from analyses of connectivity. *Neuropsychologia*, *45*(2), 209–228. https://doi.org/10.1016/j.neuropsychologia.2006.07.002

Stephan, K. E., Penny, W. D., Marshall, J. C., Fink, G. R., & Friston, K. J. (2005). Investigating the functional role of callosal connections with dynamic causal models. In *White Matter in Cognitive Neuroscience: Advances in Diffusion Tensor Imaging and Its Applications* (Vol. 1064, pp. 16–36). Annals of the New York Academy of Sciences. https://doi.org/10.1196/annals.1340.008

Stippich, C., Mohammed, J., Kress, B., Hahnel, S., Gunther, J., Konrad, F., & Sartor, K. (2003). Robust localization and lateralization of human language function: an optimized clinical functional magnetic resonance imaging protocol. *Neuroscience Letters*, *346*(1–2), 109–113. https://doi.org/10.1016/S0304-3940(03)00561-5

Stroobant, N., Buijs, D., & Vingerhoets, G. (2009). Variation in brain lateralization during various language tasks: A functional transcranial Doppler study. *Behavioural Brain Research*, *199*(2), 190–196. https://doi.org/http://dx.doi.org/10.1016/j.bbr.2008.11.040

Suarez, R. O., Whalen, S., O'Shea, J. P., & Golby, A. J. (2008). A Surgical Planning Method for Functional MRI Assessment of Language Dominance: Influences from Threshold, Region-of-Interest, and Stimulus Mode. *Brain Imaging and Behavior*, *2*(2), 59–73. https://doi.org/10.1007/s11682-007-9018-8

Sugden, D. A. (2006). Development Coordination Disorder as a Specific Learning Difficulty. Leeds Consensus Statement (2006). In *ESRC Research Seminar Series*.

Swanson, H. L., & Hsieh, C. J. (2009). Reading Disabilities in Adults: A Selective Meta-Analysis of the Literature. *Review of Educational Research*, *79*(4), 1362–1390. https://doi.org/10.3102/0034654309350931

Szaflarski, J. P., Holland, S. K., Jacola, L. M., Lindsell, C., Privitera, M. D., & Szaflarski, M. (2008). Comprehensive presurgical functional MRI language

evaluation in adult patients with epilepsy. *Epilepsy & Behavior*, *12*(1), 74–83. https://doi.org/https://doi.org/10.1016/j.yebeh.2007.07.015

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., … Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–250. https://doi.org/10.1016/J.JML.2004.03.002

Tailby, C., Abbott, D. F., & Jackson, G. D. (2017). The diminishing dominance of the dominant hemisphere: Language fMRI in focal epilepsy. *NeuroImage: Clinical*, *14*, 141–150. https://doi.org/10.1016/j.nicl.2017.01.011

Tailby, C., Weintrob, D. L., Saling, M. M., Fitzgerald, C., & Jackson, G. D. (2014). Reading difficulty is associated with failure to lateralize temporooccipital function. *Epilepsia*, *55*(5), 746–753. https://doi.org/10.1111/epi.12607

Thivard, L., Hombrouck, J., du Montcel, S. T., Delmaire, C., Cohen, L., Samson, S., … Lehericy, S. (2005). Productive and perceptive language reorganization in temporal lobe epilepsy. *NeuroImage*, *24*(3), 841–851. https://doi.org/10.1016/j.neuroimage.2004.10.001

Tops, W., Callens, M., Lammertyn, J., van Hees, V., & Brysbaert, M. (2012). Identifying students with dyslexia in higher education. *Annals of Dyslexia*, *62*(3), 186–203. https://doi.org/10.1007/s11881-012-0072-6

Torgersen, J. (1950). Situs inversus, asymmetry, and twinning. *American Journal of Human Genetics*, *2*(4), 361–370. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/14837905

Torgesen, J., Wagner, R., & Rashotte, C. (1999). *Test of Word Reading Efficiency (TOWRE)*. New York: Pearson Assessment.

Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., … Stamatakis, E. A. (2011). Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. *Brain*, *134*(2), 415–431. https://doi.org/10.1093/brain/awq369

Tzourio-Mazoyer, N., Crivello, F., & Mazoyer, B. (2018). Is the planum temporale surface area a marker of hemispheric or regional language lateralization? *Brain Structure and Function*, *223*(3), 1217–1228. https://doi.org/10.1007/s00429-017-1551-7

Tzourio-Mazoyer, N., Joliot, M., Marie, D., & Mazoyer, B. (2016). Variation in homotopic areas' activity and inter-hemispheric intrinsic connectivity with type of language lateralization: an FMRI study of covert sentence generation in 297 healthy volunteers. *Brain Structure and Function*, *221*(5), 2735–2753. https://doi.org/10.1007/s00429-015-1068-x

Tzourio-Mazoyer, N., Marie, D., Zago, L., Jobard, G., Perchey, G., Leroux, G., … Mazoyer, B. (2015). Heschl's gyrification pattern is related to speech-listening hemispheric lateralization: FMRI investigation in 281 healthy volunteers. *Brain Structure & Function*, *220*(3), 1585–1599. https://doi.org/10.1007/s00429-014-0746-4

Vallortigara, G. (2006). The evolutionary psychology of left and right: Costs and benefits of lateralization. *Developmental Psychobiology*, *48*(6), 418–427. https://doi.org/10.1002/dev.20166

van der Haegen, L., Cai, Q., & Brysbaert, M. (2012). Colateralization of Broca's area and the visual word form area in left-handers: fMRI evidence. *Brain and Language*, *122*(3), 171–178. https://doi.org/10.1016/j.bandl.2011.11.004

van der Haegen, L., Cai, Q., Seurinck, R., & Brysbaert, M. (2011). Further fMRI validation of the visual half field technique as an indicator of language laterality: A large-group analysis. *Neuropsychologia*, *49*(10), 2879–2888. https://doi.org/10.1016/j.neuropsychologia.2011.06.014

van der Knaap, L. J., & van der Ham, I. J. M. (2011). How does the corpus callosum mediate interhemispheric transfer? A review. *Behavioural Brain Research*, *223*(1), 211–221. https://doi.org/10.1016/J.BBR.2011.04.018

van der Mark, S., Klaver, P., Bucher, K., Maurer, U., Schulz, E., Brem, S., … Brandeis, D. (2011). The left occipitotemporal system in reading: Disruption of focal fMRI connectivity to left inferior frontal and inferior parietal language areas in children with dyslexia. *NeuroImage*, *54*(3), 2426–2436. https://doi.org/10.1016/j.neuroimage.2010.10.002

van Ettinger-Veenstra, H. M., Ragnehed, M., Hällgren, M., Karlsson, T., Landtblom, A.-M., Lundberg, P., & Engström, M. (2010). Right-hemispheric brain activation correlates to language performance. *NeuroImage*, *49*(4), 3481–3488. https://doi.org/10.1016/J.NEUROIMAGE.2009.10.041

van Oers, C. A. M. M., Vink, M., van Zandvoort, M. J. E., van der Worp, H. B., de Haan, E. H. F., Kappelle, L. J., … Dijkhuizen, R. M. (2010). Contribution of the left and right inferior frontal gyrus in recovery from aphasia. A functional MRI study in stroke patients with preserved hemodynamic responsiveness. *NeuroImage*, *49*(1), 885–893. https://doi.org/10.1016/j.neuroimage.2009.08.057

van Rijn, S., Aleman, A., Swaab, H., Vink, M., Sommer, I., & Kahn, R. S. (2008). Effects of an extra X chromosome on language lateralization: An fMRI study with Klinefelter men (47,XXY). *Schizophrenia Research*, *101*(1–3), 17–25. https://doi.org/10.1016/j.schres.2008.02.001

van Veelen, N. M. J., Vink, M., Ramsey, N. F., Sommer, I. E. C., van Buuren, M., Hoogendam, J. M., & Kahn, R. S. (2011). Reduced language lateralization in first-episode medication-naive schizophrenia. *Schizophrenia Research*, *127*(1–3), 195–201. https://doi.org/10.1016/j.schres.2010.12.013

Vandermosten, M., Boets, B., Wouters, J., & Ghesquière, P. (2012). A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. *Neuroscience & Biobehavioral Reviews*, *36*(6), 1532–1552. https://doi.org/10.1016/J.NEUBIOREV.2012.04.002

Vassal, F., Schneider, F., Boutet, C., Jean, B., Sontheimer, A., & Lemaire, J.-J. (2016). Combined DTI Tractography and Functional MRI Study of the

Language Connectome in Healthy Volunteers: Extensive Mapping of White Matter Fascicles and Cortical Activations. *PloS ONE*, *11*(3), e0152614. https://doi.org/10.1371/journal.pone.0152614

Vernooij, M. W., Smits, M., Wielopolski, P. A., Houston, G. C., Krestin, G. P., & van der Lugt, A. (2007). Fiber density asymmetry of the arcuate fasciculus in relation to functional hemispheric language lateralization in both right- and left-handed healthy subjects: A combined fMRI and DTI study. *NeuroImage*, *35*(3), 1064–1076. https://doi.org/https://doi.org/10.1016/j.neuroimage.2006.12.041

Vigneau, M., Beaucousin, V., Hervé, P. Y., Jobard, G., Petit, L., Crivello, F., … Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. *NeuroImage*, *54*(1), 577–593. https://doi.org/10.1016/j.neuroimage.2010.07.036

Vikingstad, E. M., George, K. P., Johnson, A. F., & Cao, Y. (2000). Cortical language lateralization in right handed normal subjects using functional magnetic resonance imaging. *Journal of the Neurological Sciences*, *175*(1), 17–27. https://doi.org/10.1016/S0022-510X(00)00269-0

Vingerhoets, G., Alderweireldt, A.-S., Vandemaele, P., Cai, Q., der Haegen, L. Van, Brysbaert, M., & Achten, E. (2013). Praxis and language are linked: Evidence from co-lateralization in individuals, with atypical language dominance. *Cortex*, *49*(1), 172–183. https://doi.org/10.1016/j.cortex.2011.11.003

Voyer, D. (1998). On the reliability and validity of noninvasive laterality measures. *Brain and Cognition*, *36*(2), 209–236. https://doi.org/10.1006/brcg.1997.0953

Waldie, K. E., Haigh, C. E., Badzakova-Trajkov, G., Buckley, J., & Kirk, I. J. (2013). Reading the wrong way with the right hemisphere. *Brain Sciences*, *3*(3), 1060–1075. https://doi.org/10.3390/brainsci3031060

Wang, D., Buckner, R. L., & Liu, H. (2014). Functional Specialization in the Human Brain Estimated By Intrinsic Hemispheric Interaction. *Journal of Neuroscience*, *34*(37), 12341–12352. https://doi.org/10.1523/jneurosci.0787-14.2014

Warmington, M., Stothard, S. E., & Snowling, M. J. (2013). Assessing dyslexia in higher education: The York adult assessment battery-revised. *Journal of Research in Special Educational Needs*, *13*(1), 48–56. https://doi.org/10.1111/j.1471-3802.2012.01264.x

Warrington, E. K., & McCarthy, R. A. (1983). Category specific access dysphasia. *Brain*, *106*(4), 859. https://doi.org/https://ezproxy-prd.bodleian.ox.ac.uk:4563/10.1093/brain/106.4.859

Warrington, E. K., & McCarthy, R. A. (1987). Categories of Knowledge - further Fractionations and an Attempted Integration. *Brain*, *110*(5), 1273–1296. https://doi.org/10.1093/brain/110.5.1273

Warrington, E. K., & Shallice, T. (1984). Category Specific Semantic

Impairments. *Brain*, *107*(3), 829–853. https://doi.org/10.1093/brain/107.3.829

Wechsler, D. (1999). *Wechsler abbreviated scale of intelligence*. New York: Psychological Corporation.

Wechsler, D. (2008). *Wechsler Adult Intelligence Test - Fourth Edition*. Texas: Pearson.

Welcome, S. E., & Joanisse, M. F. (2014). Individual differences in white matter anatomy predict dissociable components of reading skill in adults. *NeuroImage*, *96*, 261–275. https://doi.org/10.1016/j.neuroimage.2014.03.069

Westerhausen, R., Kreuder, F., Santos Sequeira, S. Dos, Walter, C., Woerner, W., Wittling, R. A., … Wittling, W. (2006). The association of macro- and microstructure of the corpus callosum and language lateralisation. *Brain and Language*, *97*(1), 80–90. https://doi.org/10.1016/J.BANDL.2005.07.133

Westerhausen, R., Kreuder, F., Woerner, W., Huster, R. J., Smit, C. M., Schweiger, E., & Wittling, W. (2006). Interhemispheric transfer time and structural properties of the corpus callosum. *Neuroscience Letters*, *409*(2), 140–145. https://doi.org/10.1016/J.NEULET.2006.09.028

Whitehouse, A. J. O., & Bishop, D. V. M. (2008). Cerebral dominance for language function in adults with specific language impairment or autism. *Brain*, *131*(12), 3193–3200. https://doi.org/10.1093/brain/awn266

Whitford, T. J., Kubicki, M., Ghorashi, S., Schneiderman, J. S., Hawley, K. J., McCarley, R. W., … Spencer, K. M. (2011). Predicting inter-hemispheric transfer time from the diffusion properties of the corpus callosum in healthy individuals and schizophrenia patients: A combined ERP and DTI study. *NeuroImage*, *54*(3), 2318–2329. https://doi.org/10.1016/j.neuroimage.2010.10.048

Wilke, M., & Lidzba, K. (2007). LI-tool: A new toolbox to assess lateralization in functional MR-data. *Journal of Neuroscience Methods*, *163*(1), 128–136. https://doi.org/10.1016/j.jneumeth.2007.01.026

Wilke, M., & Schmithorst, V. J. (2006). A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data. *NeuroImage*, *33*(2), 522–530. https://doi.org/https://doi.org/10.1016/j.neuroimage.2006.07.010

Wilson, A. C., & Bishop, D. V. M. (2018). Resounding failure to replicate links between developmental language disorder and cerebral lateralisation. *PeerJ*, *6*, e4217. https://doi.org/10.7717/peerj.4217

Wiseheart, R., & Altmann, L. J. P. (2018). Spoken sentence production in college students with dyslexia: working memory and vocabulary effects. *International Journal of Language & Communication Disorders*, *53*(2), 355–369. https://doi.org/10.1111/1460-6984.12353

Wiseheart, R., Altmann, L., Park, H., & Lombardino, L. (2009). Sentence comprehension in young adults with developmental dyslexia. *Annals of*

Dyslexia, *59*(2), 151–167. https://doi.org/10.1007/s11881-009-0028-7

Wood, A. G., Harvey, A. S., Wellard, R. M., Abbott, D. F., Anderson, V., Kean, M., … Jackson, G. D. (2004). Language cortex activation in normal children. *Neurology*, *63*(6), 1035–1044. https://doi.org/https://doi.org/10.1212/01.WNL.0000140707.61952.CA

Woodhead, Z., Bradshaw, A., Wilson, A., Thompson, P., & Bishop, D. (2019). Testing the unitary theory of language lateralization using functional transcranial Doppler sonography in adults. *Royal Society Open Science*, *6*(3), 181801. https://doi.org/10.1098/rsos.181801

Wright, P., Stamatakis, E. A., & Tyler, L. K. (2012). Differentiating Hemispheric Contributions to Syntax and Semantics in Patients with Left-Hemisphere Lesions. *Journal of Neuroscience*, *32*(24), 8149–8157. https://doi.org/10.1523/JNEUROSCI.0485-12.2012

Xu, M., Yang, J., Siok, W. T., & Tan, L. H. (2015). Atypical lateralization of phonological working memory in developmental dyslexia. *Journal of Neurolinguistics*, *33*, 67–77. https://doi.org/10.1016/J.JNEUROLING.2014.07.004

Yamada, Y., Stevens, C., Dow, M., Harn, B. A., Chard, D. J., & Neville, H. J. (2011). Emergence of the neural network for reading in five-year-old beginning readers of different levels of pre-literacy abilities: An fMRI study. *NeuroImage*, *57*(3), 704–713. https://doi.org/10.1016/J.NEUROIMAGE.2010.10.057

Zaca, D., Jarso, S., & Pillai, J. J. (2013). Role of Semantic Paradigms for Optimization of Language Mapping in Clinical fMRI Studies. *American Journal of Neuroradiology*, *34*(10), 1966–1971. https://doi.org/10.3174/ajnr.A3628

Zaidel, E., Clarke, J., & Suyenobu, B. (1990). Hemispheric Independence: a paradigm case for cognitive neuroscience. In *Neurobiology of Higher Cognitive Functions* (p. 59). New York: Guilford Press.

Zhao, J., Thiebaut de Schotten, M., Altarelli, I., Dubois, J., & Ramus, F. (2016). Altered hemispheric lateralization of white matter pathways in developmental dyslexia: Evidence from spherical deconvolution tractography. *Cortex*, *76*, 51–62. https://doi.org/10.1016/J.CORTEX.2015.12.004

Appendix 1- Systematic literature search

List of 76 selected studies for inclusion in systematic reviews presented in Chapters 1 and 2

1. Abbott, D. F., Waites, A. B., Lillywhite, L. M., & Jackson, G. D. (2010). fMRI assessment of language lateralization: An objective approach. *Neuroimage, 50*(4), 1446-1455. doi:10.1016/j.neuroimage.2010.01.059
2. Adcock, J., Wise, R., Oxbury, J., Oxbury, S., & Matthews, P. (2003). Quantitative fMRI assessment of the differences in lateralization of language-related brain activation in patients with temporal lobe epilepsy. *Neuroimage, 18*(2), 423-438. doi:10.1016/S1053-8119(02)00013-7
3. Allendorfer, J. B., Hernando, K. A., Hossain, S., Nenert, R., Holland, S. K., & Szaflarski, J. P. (2016). Arcuate fasciculus asymmetry has a hand in language function but not handedness. *Human Brain Mapping, 37*(9), 3297-3309. doi:10.1002/hbm.23241
4. Baciu, M., Juphard, A., Cousin, E., & Le Bas, J. (2005). Evaluating fMRI methods for assessing hemispheric language dominance in healthy subjects. *European Journal of Radiology, 55*(2), 209-218. doi:10.1016/j.ejrad.2004.11.004
5. Backes, W., Deblaere, K., Vonck, K., Kessels, A., Boon, P., Hofman, P., et al. (2005). Language activation distributions revealed by fMRI in post-operative epilepsy patients: Differences between left- and right-sided resections. *Epilepsy Research, 66*(1-3), 1-12. doi:10.1016/j.eplepsyres.2005.06.007
6. Berl, M. M., Zimmaro, L. A., Khan, O. I., Dustin, I., Ritzl, E., Duke, E. S., et al. (2014). Characterization of atypical language activation patterns in focal epilepsy. *Annals of Neurology, 75*(1), 33-42. doi:10.1002/ana.24015
7. Bethmann, A., Tempelmann, C., De Bleser, R., Scheich, H., & Brechmann, A. (2007). Determining language laterality by fMRI and dichotic listening. *Brain Research, 1133*(1), 145-157. doi:10.1016/j.brainres.2006.11.057
8. Binder, J. R., Swanson, S. J., Hammeke, T. A., & Sabsevitz, D. S. (2008). A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia, 49*(12), 1980-1997. doi:10.1111/j.1528-1167.2008.01683.x
9. Brennan, N. M. P., Whalen, S., Branco, D. d. M., O'Shea, J. P., Norton, I. H., & Golby, A. J. (2007). Object naming is a more sensitive measure of speech localization than number counting: Converging evidence from direct cortical stimulation and fMRI. *Neuroimage, 37*, S100-S108. doi:10.1016/j.neuroimage.2007.04.052
10. Clements, A. M., Rimrodt, S. L., Abel, J. R., Blankner, J. G., Mostofsky, S. H., Pekar, J. J., et al. (2006). Sex differences in cerebral laterality of language and visuospatial processing. *Brain and Language, 98*(2), 150-158. doi:10.1016/j.bandl.2006.04.007
11. Cousin, E., Peyrin, C., Pichat, C., Lamalle, L., Le Bas, J., & Baciu, M. (2007). Functional MRI approach for assessing hemispheric

predominance of regions activated by a phonological and a semantic task. *European Journal of Radiology, 63*(2), 274-285. doi:10.1016/j.ejrad.2007.01.030
12. Deblaere, K., Backes, W., Hofman, P., Vandemaele, P., Boon, P., Vonck, K., et al. (2002). Developing a comprehensive presurgical functional MRI protocol for patients with intractable temporal lobe epilepsy: A pilot study. *Neuroradiology, 44*(8), 667-673. doi:10.1007/s00234-002-0800-4
13. Dodoo-Schittko, F., Rosengarth, K., Doenitz, C., & Greenlee, M. W. (2012). Assessing language dominance with functional MRI: The role of control tasks and statistical analysis. *Neuropsychologia, 50*(11), 2684-2691. doi:10.1016/j.neuropsychologia.2012.07.032
14. Doucet, G. E., Pustina, D., Skidmore, C., Sharan, A., Sperling, M. R., & Tracy, J. I. (2015). Resting-state functional connectivity predicts the strength of hemispheric lateralization for language processing in temporal lobe epilepsy and normals. *Human Brain Mapping, 36*(1), 288-303. doi:10.1002/hbm.22628
15. Drager, B., Jansen, A., Bruchmann, S., Forster, A., Pleger, B., P'Zwitserlood, et al. (2004). How does the brain accommodate to increased task difficulty in word finding? A functional MRI study. *Neuroimage, 23*(3), 1152-1160. doi:10.1016/j.neuroimage.2004.07.005
16. Fernandez, G., de Greiff, A., von Oertzen, J., Reuber, M., Lun, S., Klaver, P., et al. (2001). Language mapping in less than 15 minutes: Real-time functional MRI during routine clinical investigation. *Neuroimage, 14*(3), 585-594. doi:10.1006/nimg.2001.0854
17. Fesl, G., Bruhns, P., Rau, S., Wiesmann, M., Ilmberger, J., Kegel, G., et al. (2010). Sensitivity and reliability of language laterality assessment with a free reversed association task-a fMRI study. *European Radiology, 20*(3), 683-695. doi:10.1007/s00330-009-1602-4
18. Gaillard, W., Balsamo, L., Xu, B., Grandin, C., Braniecki, S., Papero, P., et al. (2002). Language dominance in partial epilepsy patients identified with an fMRI reading task. *Neurology, 59*(2), 256-265.
19. Gaillard, W., Sachs, B., Whitnah, J., Ahmad, Z., Balsamo, L., Petrella, J., et al. (2003). Developmental aspects of language processing: FMRI of verbal fluency in children and adults. *Human Brain Mapping, 18*(3), 176-185. doi:10.1002/hbm.10091
20. Haeberling, I. S., Badzakova-Trajkov, G., & Corballis, M. C. (2011). Callosal tracts and patterns of hemispheric dominance: A combined fMRI and DTI study. *Neuroimage, 54*(2), 779-786. doi:10.1016/j.neuroimage.2010.09.072
21. Haeberling, I. S., Steinemann, A., & Corballis, M. C. (2016). Cerebral asymmetry for language: Comparing production with comprehension. *Neuropsychologia, 80*, 17-23. doi:10.1016/j.neuropsychologia.2015.11.002
22. Harrington, G., Buonocore, M., & Farias, S. (2006). Intrasubject reproducibility of functional MR imaging activation in language tasks. *American Journal of Neuroradiology, 27*(4), 938-944.
23. Hernandez, N., Andersson, F., Edjlali, M., Hommet, C., Cottier, J. P., Destrieux, C., et al. (2013). Cerebral functional asymmetry and phonological performance in dyslexic adults. *Psychophysiology, 50*(12), 1226-1238. doi:10.1111/psyp.12141

24. Hund-Georgiadis, M., Lex, U., Friederici, A., & von Cramon, D. (2002). Non-invasive regime for language lateralization in right- and left-handers by means of functional MRI and dichotic listening. *Experimental Brain Research, 145*(2), 166-176. doi:10.1007/s00221-002-1090-0
25. Hund-Georgiadis, M., Lex, U., & von Cramon, D. (2001). Language dominance assessment by means of fMRI: Contributions from task design, performance, and stimulus modality. *Journal of Magnetic Resonance Imaging, 13*(5), 668-675. doi:10.1002/jmri.1094
26. Hunter, Z. R., & Brysbaert, M. (2008). Visual half-field experiments are a good measure of cerebral language dominance if used properly: Evidence from fMR1. *Neuropsychologia, 46*(1), 316-325. doi:10.1016/j.neuropsychologia.2007.07.007
27. Jansen, A., Menke, R., Sommer, J., Foerster, A. F., Bruchmann, S., Hempleman, J., et al. (2006). The assessment of hemispheric lateralization in functional MRI - robustness and reproducibility. *Neuroimage, 33*(1), 204-217. doi:10.1016/j.neuroimage.2006.06.019
28. Jensen-Kondering, U. R., Ghobadi, Z., Wolff, S., Jansen, O., & Ulmer, S. (2012). Acoustically presented semantic decision-making tasks provide a robust depiction of the temporo-parietal speech areas. *Journal of Clinical Neuroscience, 19*(3), 428-433. doi:10.1016/j.jocn.2011.04.038
29. Kennan, R., Kim, D., Maki, A., Koizumi, H., & Constable, R. (2002). Non-invasive assessment of language lateralization by transcranial near infrared optical topography and functional MRI. *Human Brain Mapping, 16*(3), 183-189. doi:10.1002/hbm.10039
30. Kleinhans, N. M., Mueller, R., Cohen, D. N., & Courchesne, E. (2008). Atypical functional lateralization of language in autism spectrum disorders. *Brain Research, 1221*, 115-125. doi:10.1016/j.brainres.2008.04.080
31. Knecht, S., Jansen, A., Frank, A., van Randenborgh, J., Sommer, J., Kanowski, A., et al. (2003). How atypical is atypical language dominance? *Neuroimage, 18*(4), 917-927. doi:10.1016/S1053-8119(03)00039-9
32. Krainik, A., Lehéricy, S., Duffau, H., Capelle, L., Chainay, H., Cornu, P., et al. (2003). Postoperative speech disorder after medial frontal surgery: Role of the supplementary motor area. *Neurology, 60*(4), 587-594. doi:10.1212/01.WNL.0000048206.07837.59
33. Lohmann, H., Deppe, M., Jansen, A., Schwindt, T., & Knecht, S. (2004). Task repetition can affect functional magnetic resonance imaging-based measures of language lateralization and lead to pseudoincreases in bilaterality. *Journal of Cerebral Blood Flow and Metabolism, 24*(2), 179-187. doi:10.1097/01.WCb.0000100066.36077.91
34. Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Perchey, G., et al. (2014). Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *Plos One, 9*(6), e101165. doi:10.1371/journal.pone.0101165
35. Miro, J., Ripolles, P., Lopez-Barroso, D., Vila-Ballo, A., Juncadella, M., de Diego-Balaguer, R., et al. (2014). Atypical language organization in temporal lobe epilepsy revealed by a passive semantic paradigm. *Bmc Neurology, 14*, 98. doi:10.1186/1471-2377-14-98

36. Morrison, M. A., Churchill, N. W., Cusimano, M. D., Schweizer, T. A., Das, S., & Graham, S. J. (2016). Reliability of task-based fMRI for preoperative planning: A test-retest study in brain tumor patients and healthy controls. *Plos One, 11*(2), e0149547. doi:10.1371/journal.pone.0149547
37. Nadkarni, T. N., Andreoli, M. J., Nair, V. A., Yin, P., Young, B. M., Kundu, B., et al. (2015). Usage of fMRI for pre-surgical planning in brain tumor and vascular lesion patients: Task and statistical threshold effects on language lateralization. *Neuroimage-Clinical, 7*, 415-423. doi:10.1016/j.nicl.2014.12.014
38. Niskanen, E., Kononen, M., Villberg, V., Nissi, M., Ranta-aho, P., Saisanen, L., et al. (2012). The effect of fMRI task combinations on determining the hemispheric dominance of language functions. *Neuroradiology, 54*(4), 393-405. doi:10.1007/s00234-011-0959-7
39. Ocklenburg, S., Hugdahl, K., & Westerhausen, R. (2013). Structural white matter asymmetries in relation to functional asymmetries during speech perception and production. *Neuroimage, 83*, 1088-1097. doi:10.1016/j.neuroimage.2013.07.076
40. Orellana, C. M., Visch-Brink, E., Vernooij, M., Kalloe, S., Satoer, D., Vincent, A., et al. (2015). Crossed cerebrocerebellar language lateralization: An additional diagnostic feature for assessing atypical language representation in presurgical functional MR imaging. *American Journal of Neuroradiology, 36*(3), 518-524. doi:10.3174/ajnr.A4147
41. Partovi, S., Jacobi, B., Rapps, N., Zipp, L., Karimi, S., Rengier, F., et al. (2012). Clinical standardized fMRI reveals altered language lateralization in patients with brain tumor. *American Journal of Neuroradiology, 33*(11), 2151-2157. doi:10.3174/ajnr.A3137
42. Partovi, S., Konrad, F., Karimi, S., Rengier, F., Lyo, J. K., Zipp, L., et al. (2012). Effects of covert and overt paradigms in clinical language fMRI. *Academic Radiology, 19*(5), 518-525. doi:10.1016/j.acra.2011.12.017
43. Perlaki, G., Horvath, R., Orsi, G., Aradi, M., Auer, T., Varga, E., et al. (2013). White-matter microstructure and language lateralization in left-handers: A whole-brain MRI analysis. *Brain and Cognition, 82*(3), 319-328. doi:10.1016/j.bandc.2013.05.005
44. Pravata, E., Sestieri, C., Mantini, D., Briganti, C., Colicchio, G., Marra, C., et al. (2011). Functional connectivity MR imaging of the language network in patients with drug-resistant epilepsy. *American Journal of Neuroradiology, 32*(3), 532-540. doi:10.3174/ajnr.A2311
45. Propper, R. E., O'Donnell, L. J., Whalen, S., Tie, Y., Norton, I. H., Suarez, R. O., et al. (2010). A combined fMRI and DTI examination of functional language lateralization and arcuate fasciculus structure: Effects of degree versus direction of hand preference. *Brain and Cognition, 73*(2), 85-92. doi:10.1016/j.bandc.2010.03.004
46. Ramsey, N., Sommer, I., Rutten, G., & Kahn, R. (2001). Combined analysis of language tasks in fMRI improves assessment of hemispheric dominance for language functions in individual subjects. *Neuroimage, 13*(4), 719-733. doi:10.1006/nimg.2000.0722
47. Razafimandimby, A., Maiza, O., Herve, P., Lecardeur, L., Delamillieure, P., Brazo, P., et al. (2007). Stability of functional language lateralization

over time in schizophrenia patients. *Schizophrenia Research, 94*(1-3), 197-206. doi:10.1016/j.schres.2007.04.011
48. Ruff, I. M., Brennan, N. M. P., Peck, K. K., Hou, B. L., Tabar, V., Brennan, C. W., et al. (2008). Assessment of the language laterality index in patients with brain tumor using functional MR imaging: Effects of thresholding, task selection, and prior surgery. *American Journal of Neuroradiology, 29*(3), 528-535. doi:10.3174/ajnr.A0841
49. Rutten, G., Ramsey, N., van Rijen, P., & van Veelen, C. (2002). Reproducibility of fMRI-determined language lateralization in individual subjects. *Brain and Language, 80*(3), 421-437. doi:10.1006/brln.2001.2600
50. Sanjuan, A., Bustamante, J., Forn, C., Ventura-Campos, N., Barros-Loscertales, A., Martinez, J., et al. (2010). Comparison of two fMRI tasks for the evaluation of the expressive language function. *Neuroradiology, 52*(5), 407-415. doi:10.1007/s00234-010-0667-8
51. Sanjuan, A., Forn, C., Ventura-Campos, N., Rodriguez-Pujadas, A., Garcia-Porcar, M., Belloch, V., et al. (2010). The sentence verification task: A reliable fMRI protocol for mapping receptive language in individual subjects. *European Radiology, 20*(10), 2432-2438. doi:10.1007/s00330-010-1814-7
52. Seghier, M., Lazeyras, F., Pegna, A., Annoni, J., Zimine, I., Mayer, E., et al. (2004). Variability of fMRI activation during a phonological and semantic language task in healthy subjects. *Human Brain Mapping, 23*(3), 140-155. doi:10.1002/hbm.20053
53. Seghier, M. L., Kherif, F., Josse, G., & Price, C. J. (2011). Regional and hemispheric determinants of language laterality: Implications for preoperative fMRI. *Human Brain Mapping, 32*(10), 1602-1614. doi:10.1002/hbm.21130
54. Sepeta, L. N., Berl, M. M., Wilke, M., You, X., Mehta, M., Xu, B., et al. (2016). Age-dependent mesial temporal lobe lateralization in language fMRI. *Epilepsia, 57*(1), 122-130. doi:10.1111/epi.13258
55. Somers, M., Neggers, S. F., Diederen, K. M., Boks, M. P., Kahn, R. S., & Sommer, I. E. (2011). The measurement of language lateralization with functional transcranial doppler and functional MRI: A critical evaluation. *Frontiers in Human Neuroscience, 5*, 31. doi:10.3389/fnhum.2011.00031
56. Sommer, I., Ramsey, N., Mandl, R., & Kahn, R. (2003). Language lateralization in female patients with schizophrenia: An fMRI study. *Schizophrenia Research, 60*(2-3), 183-190. doi:10.1016/S0920-9964(02)00300-6
57. Stippich, C., Mohammed, J., Kress, B., Hahnel, S., Gunther, J., Konrad, F., et al. (2003). Robust localization and lateralization of human language function: An optimized clinical functional magnetic resonance imaging protocol. *Neuroscience Letters, 346*(1-2), 109-113. doi:10.1016/S0304-3940(03)00561-5
58. Suarez, R. O., Whalen, S., O'Shea, J. P., & Golby, A. J. (2008). A surgical planning method for functional MRI assessment of language dominance: Influences from threshold, region-of-interest, and stimulus mode. *Brain Imaging and Behavior, 2*(2), 59-73. doi:10.1007/s11682-007-9018-8

59. Sveller, C., Briellmann, R. S., Saling, M. M., Lillywhite, L., Abbott, D. F., Masterton, R. A. J., et al. (2006). Relationship between language lateralization and handedness in left-hemispheric partial epilepsy. *Neurology, 67*(10), 1813-1817. doi:10.1212/01.wnl.0000244465.74707.42
60. Szaflarski, J. P., Holland, S. K., Jacola, L. M., Lindsell, C., Privitera, M. D., & Szaflarski, M. (2008). Comprehensive presurgical functional MRI language evaluation in adult patients with epilepsy. *Epilepsy & Behavior, 12*(1), 74-83. doi:10.1016/j.yebeh.2007.07.015
61. Tailby, C., Weintrob, D. L., Saling, M. M., Fitzgerald, C., & Jackson, G. D. (2014). Reading difficulty is associated with failure to lateralize temporooccipital function. *Epilepsia, 55*(5), 746-753. doi:10.1111/epi.12607
62. Thivard, L., Hombrouck, J., du Montcel, S., Delmaire, C., Cohen, L., Samson, S., et al. (2005). Productive and perceptive language reorganization in temporal lobe epilepsy. *Neuroimage, 24*(3), 841-851. doi:10.1016/j.neuroimage.2004.10.001
63. Tie, Y., Suarez, R. O., Whalen, S., Radmanesh, A., Norton, I. H., & Golby, A. J. (2009). Comparison of blocked and event-related fMRI designs for pre-surgical language mapping. *Neuroimage, 47*, T107-T115. doi:10.1016/j.neuroimage.2008.11.020
64. Tzourio-Mazoyer, N., Marie, D., Zago, L., Jobard, G., Perchey, G., Leroux, G., et al. (2015). Heschl's gyrification pattern is related to speech-listening hemispheric lateralization: FMRI investigation in 281 healthy volunteers. *Brain Structure & Function, 220*(3), 1585-1599. doi:10.1007/s00429-014-0746-4
65. Tzourio-Mazoyer, N., Joliot, M., Marie, D., & Mazoyer, B. (2016). Variation in homotopic areas' activity and inter-hemispheric intrinsic connectivity with type of language lateralization: An FMRI study of covert sentence generation in 297 healthy volunteers. *Brain Structure & Function, 221*(5), 2735-2753. doi:10.1007/s00429-015-1068-x
66. Van der Haegen, L., Cai, Q., & Brysbaert, M. (2012). Colateralization of broca's area and the visual word form area in left-handers: FMRI evidence. *Brain and Language, 122*(3), 171-178. doi:10.1016/j.bandl.2011.11.004
67. Van der Haegen, L., Cai, Q., Seurinck, R., & Brysbaert, M. (2011). Further fMRI validation of the visual half field technique as an indicator of language laterality: A large-group analysis. *Neuropsychologia, 49*(10), 2879-2888. doi:10.1016/j.neuropsychologia.2011.06.014
68. van Oers, C. A. M. M., Vink, M., van Zandvoort, M. J. E., van der Worp, H. B., de Haan, E. H. F., Kappelle, L. J., et al. (2010). Contribution of the left and right inferior frontal gyrus in recovery from aphasia. A functional MRI study in stroke patients with preserved hemodynamic responsiveness. *Neuroimage, 49*(1), 885-893. doi:10.1016/j.neuroimage.2009.08.057
69. van Rijn, S., Aleman, A., Swaab, H., Vink, M., Sommer, I., & Kahn, R. S. (2008). Effects of an extra X chromosome on language lateralization: An fMRI study with klinefelter men (47,XXY). *Schizophrenia Research, 101*(1-3), 17-25. doi:10.1016/j.schres.2008.02.001

70. van Veelen, N. M. J., Vink, M., Ramsey, N. F., Sommer, I. E. C., van Buuren, M., Hoogendam, J. M., et al. (2011). Reduced language lateralization in first-episode medication-naive schizophrenia. *Schizophrenia Research, 127*(1-3), 195-201. doi:10.1016/j.schres.2010.12.013
71. Vassal, F., Schneider, F., Boutet, C., Jean, B., Sontheimer, A., & Lemaire, J. (2016). Combined DTI tractography and functional MRI study of the language connectome in healthy volunteers: Extensive mapping of white matter fascicles and cortical activations. *Plos One, 11*(3), e0152614. doi:10.1371/journal.pone.0152614
72. Vernooij, M. W., Smits, M., Wielopolski, P. A., Houston, G. C., Krestin, G. P., & van der Lugt, A. (2007). Fiber density asymmetry of the arcuate fasciculus in relation to functional hemispheric language lateralization in both right- and left-handed healthy subjects: A combined fMRI and DTI study. *Neuroimage, 35*(3), 1064-1076. doi:10.1016/j.neuroimage.2006.12.041
73. Vikingstad, E., George, K., Johnson, A., & Cao, Y. (2000). Cortical language lateralization in right handed normal subjects using functional magnetic resonance imaging. *Journal of the Neurological Sciences, 175*(1), 17-27. doi:10.1016/S0022-510X(00)00269-0
74. Vingerhoets, G., Alderweireldt, A., Vandemaele, P., Cai, Q., Van der Haegen, L., Brysbaert, M., et al. (2013). Praxis and language are linked: Evidence from co-lateralization in individuals, with atypical language dominance. *Cortex, 49*(1), 172-183. doi:10.1016/j.cortex.2011.11.003
75. Wilke, M., & Lidzba, K. (2007). LI-tool: A new toolbox to assess lateralization in functional MR-data. *Journal of Neuroscience Methods, 163*(1), 128-136. doi:10.1016/j.jneumeth.2007.01.026
76. Zaca, D., Jarso, S., & Pillai, J. J. (2013). Role of semantic paradigms for optimization of language mapping in clinical fMRI studies. *American Journal of Neuroradiology, 34*(10), 1966-1971. doi:10.3174/ajnr.A3628