The Physiology of Dementia

Network reorganisation in progressive non-fluent aphasia as a model of neurodegeneration

Thomas Edmund Cope

Gonville and Caius College, University of Cambridge
Department of Clinical Neuroscience

This dissertation is submitted for the degree of

Doctor of Philosophy

January 2018
Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. Several of the chapters have been accepted for publication with co-authors. In each case a preface to the chapter sets out my personal contribution to the work:

Chapter 1 forms the basis of Chapter 25 of the upcoming Oxford Textbook of Neuropsychiatry.

Chapter 2 forms the basis of a paper that has been published in Brain (Cope et al., 2018).

Chapter 3 forms the basis of a paper that has been published in Nature Communications (Cope et al., 2017a).

Chapter 4 forms the basis of a paper that has been published in Neuropsychologia (Cope et al., 2017b).

This dissertation is within the word limit requirements set by the Clinical Medicine Degree Committee and the Board of Graduate Studies.

I confirm that no part of this dissertation has previously been submitted for any other qualification.
Acknowledgements

I am, and always will be, grateful to my supervisor, Professor James Rowe, for his support of my research and personal development. He has provided me with the perfect balance of guidance and freedom to allow me to make impactful and meaningful discoveries while acquiring my own investigative skills and analytical approach. His influence has extended far beyond the scientific, being a fantastic clinical mentor, pastoral support, inspirational role model, and good friend.

I am grateful to my advisor, Dr Matt Davis, who has been generous with his time, collaborative in his approach, and constructive in his criticism.

I am grateful to my former supervisor, Prof Tim Griffiths, who has always been willing to discuss my projects and never fails to provide insight when I lack clarity. I hope that our productive collaboration continues for many years to come.

I am grateful to Prof Karalyn Patterson, from whom I have learned everything I now know about aphasia. I have been extremely privileged to work next to you scientifically, and to share my consultations with you clinically.

I am grateful to the many others with whom I have worked during my PhD. While the mandatory declaration begins by stating that this work is entirely my own, in truth none of it could have been accomplished without the collaborations that are described in the preface to each chapter. I am especially indebted to Dr Ediz Sohoglu, who was generous in his time and computer code getting me up to speed with his experiments, on which chapter 3 builds; Dr Laura Hughes and Mr Maarten van Casteren for teaching me the practicalities of MEG acquisition; Prof Chris Petkov, whose boundless enthusiasm kept chapter 4 on the road; and Mr P Simon Jones, who taught me MRI analysis with unfailing patience.

Most of all, I am grateful to my wife, Wei, whose support and understanding knows no bounds. I don’t know how she puts up with me.

This work was primarily supported by the Patrick Berthoud Charitable Trust and the Association of British Neurologists, with additional support from the NIHR Biomedical Research Centre and the Medical Research Council.
Dedication

I dedicate this dissertation to my daughters Ellie and Catherine, who bring sunshine wherever they go.
Table of Contents

Declaration ............................................................................................................................................... ii
Acknowledgements ............................................................................................................................... iii
Dedication ................................................................................................................................................ iv
Table of Contents .................................................................................................................................. v
Table of Figures ....................................................................................................................................... x
Table of Tables ....................................................................................................................................... xii
General Introduction ............................................................................................................................. 1
Chapter 1: The current clinical context of dementia .............................................................................. 4
  1.1 Preface ............................................................................................................................................... 4
  1.2 Introduction ....................................................................................................................................... 5
  1.3 Core Progressive Dementias ............................................................................................................. 6
    1.3.1 Alzheimer’s Disease – Amnestic ................................................................................................. 6
    1.3.2 Alzheimer’s Disease – Posterior Cortical Atrophy – Benson’s Syndrome .... 10
    1.3.3 Alzheimer’s Disease – Logopenic and Mixed Primary Progressive Aphasia .12
    1.3.4 Alzheimer’s Disease – Dysexecutive/Behavioural Variant ................................. 14
    1.3.5 Frontotemporal Dementia – Behavioural Variant – bvFTD ............................. 15
    1.3.6 Frontotemporal Dementia – Semantic Dementia - SD ................................. 19
    1.3.7 Frontotemporal Dementia – non-fluent variant Primary Progressive Aphasia – nfvPPA/PNFA........................................................................................................... 22
  1.4 Dementias associated with Movement Disorders .......................................................................... 26
    1.4.1 Lewy Body related – Dementia with Lewy Bodies – DLB .............................................. 26
    1.4.2 Lewy Body related – Parkinson’s Disease Dementia - PDD ........................................... 27
    1.4.3 Progressive Supranuclear Palsy - PSP ................................................................................. 28
    1.4.4 Corticobasal Syndrome........................................................................................................... 30
    1.4.5 Multiple System Atrophy - MSA ............................................................................................ 33
    1.4.6 Huntington’s Disease - HD ..................................................................................................... 34
    1.4.7 Normal Pressure Hydrocephalus - NPH .............................................................................. 35
    1.4.8 Niemann-Pick type C .............................................................................................................. 36
    1.4.9 Spinocerebellar Ataxia - SCA ................................................................................................. 37
1.5 Cognitive impairments that may or may not progress.............................. 38
  1.5.1 Mild Cognitive Impairment............................................................... 38
  1.5.2 Subjective Cognitive Impairment....................................................... 40
  1.5.3 Transient Global Amnesia - TGA..................................................... 41
  1.5.4 Transient Epileptic Amnesia - TEA .................................................. 41
1.6 Challenges and opportunities...................................................................... 42
1.7 Chapter Summary: .................................................................................. 44

Chapter 2: Tau burden and the functional connectome in Alzheimer’s disease and
progressive supranuclear palsy........................................................................ 45
  2.1 Preface ................................................................................................. 45
  2.2 Introduction .......................................................................................... 46
  2.3 Methods................................................................................................ 48
    2.3.1 Participants....................................................................................... 48
    2.3.2 Study Procedures............................................................................. 51
    2.3.3 MRI data acquisition and pre-processing......................................... 51
    2.3.4 Functional Connectivity Assessment.............................................. 53
    2.3.5 Tau Burden Assessment................................................................. 55
    2.3.6 Statistical Approach....................................................................... 56
  2.4 Results.................................................................................................... 58
    2.4.1 Resting state connectivity differences exist between groups............ 58
    2.4.2 Findings in AD but not PSP are consistent with trans-neuronal spread of tau.............................................................................................................. 59
    2.4.3 Findings in PSP but not AD are consistent with hub vulnerability due to
        metabolic demand and lack of trophic support.................................... 66
    2.4.4 Alzheimer’s disease and PSP have opposite effects on the strength of cortical
        functional connectivity ....................................................................... 70
    2.4.5 Reorganisation of cortical functional connectivity reflects cortical vs
        subcortical pathology ........................................................................... 74
    2.4.6 Local functional connectivity reorganisation is related to local \(^{18}\text{F}\text{AV-1451}
        binding potential in Alzheimer’s disease............................................ 83
  2.5 Discussion............................................................................................... 84
    2.5.1 Insights into the mechanisms of disease progression in humans........ 85
    2.5.2 Is the difference between AD and PSP mediated by tau isoform or intrinsic
        connectivity? ....................................................................................... 87
2.5.3 Cross-sectional data reveal patterns less visible at the group level ..........88
2.5.4 Reorganisation of brain networks..................................................................89
2.5.5 Study limitations.........................................................................................90
2.6 Conclusion .......................................................................................................91
2.7 Chapter Summary: ..........................................................................................92

Chapter 3: Top-down frontal contributions to predictive processes in speech perception .................................................................94

3.1 Preface ..............................................................................................................94
3.2 Introduction .......................................................................................................97
3.3 Methods: ............................................................................................................100
  3.3.1 Participants .......................................................... ............................................100
  3.3.2 Voxel Based Morphometry ........................................................................101
  3.3.3 Modifications to Sohoglu MEG paradigm ..................................................103
  3.3.4 Behavioural data stimuli and procedure ......................................................104
  3.3.5 Behavioural data modelling ......................................................................105
  3.3.6 Alternative data models ...........................................................................107
  3.3.7 MEG and EEG data acquisition and analysis .............................................108
  3.3.8 Sensor-space evoked analysis ....................................................................110
  3.3.9 Evoked data source reconstruction ............................................................110
  3.3.10 Sensor-space induced analysis ..................................................................111
  3.3.11 Induced data source reconstruction ..........................................................112
  3.3.12 Coherence and Connectivity Analyses .....................................................112
3.4 Results: ............................................................................................................113
  3.4.1 Structural consequences of nfvPPA ...........................................................113
  3.4.2 Subjective speech perception symptoms in nfvPPA ....................................114
  3.4.3 Evoked neural responses during the reconciliation of predictions, and their cortical origins..........................................................115
  3.4.4 Behavioural Experiment 1: Vocoded Word Clarity Rating ....................125
  3.4.5 Behavioural Experiment 2: Vocoded Word Identification ......................126
  3.4.6 Bayesian modelling of experiments 1 and 2 ...........................................127
  3.4.7 Induced oscillatory dynamics ....................................................................130
  3.4.8 Coherence and Connectivity during the reconciliation of predictions ......135
3.5 Discussion: ................................................................. 137
  3.5.1 Fronto-temporal dissociations in nfvPPA .................................. 139
  3.5.2 Frontotemporal hierarchy and predictive coding ...................... 140
  3.5.3 Study limitations ..................................................... 142
3.6 Conclusion: .................................................................. 143
3.7 Chapter Summary: ........................................................ 144

Chapter 4: Artificial grammar learning in vascular and progressive non-fluent aphasias .................................................. 145

4.1 Preface .................................................................. 145
4.2 Introduction ............................................................. 147
4.3 Methods .................................................................. 150
  4.3.1 Participants ............................................................. 150
  4.3.2 Stimuli ................................................................ 155
  4.3.3 Procedure ............................................................. 158
  4.3.4 Stroke lesion mapping ............................................. 160
  4.3.5 nfvPPA atrophy mapping ....................................... 161
  4.3.6 Analysis ................................................................. 161
4.4 Results .................................................................. 165
  4.4.1 Discriminability of sequencing rules .............................. 165
  4.4.2 Bias towards permissibility ....................................... 168
  4.4.3 Implicit learning through repeated exposure-test cycles ....... 169
  4.4.4 Separation of rule learning between linguistic and non-linguistic sequences ................................................ 170
  4.4.5 Relationship to diagnostic category and disease severity ....... 171
  4.4.6 Relationship to structural anatomy .............................. 171
4.5 Discussion ................................................................. 172
  4.5.1 Rule acquisition differs when structurally identical sequences are comprised of linguistic or non-linguistic stimuli ............................................................. 172
  4.5.2 Artificial grammar learning in aphasia is similar across aetiologies ...... 174
  4.5.3 Agrammatic aphasic patients are similarly impaired for both complex and simple sequencing operations ................................................................. 175
  4.5.4 Patients with aphasia show improved performance over repeated cycles. . 176
  4.5.5 Brain behaviour relationships ................................... 178
4.5.6 Study limitations.......................................................................................................................... 179
4.6 Conclusion ....................................................................................................................................... 179
4.7 Appendix: Supplementary tables .................................................................................................. 181
4.8 Chapter Summary: .......................................................................................................................... 182

Chapter 5: Review and Synthesis........................................................................................................ 183

5.1 Network organisation and reorganisation in neurodegeneration .............................................. 183

5.2 Clinical implications of inflexible predictions in nfvPPA ............................................................ 187
  5.2.1 Inflexible predictions can explain speech comprehension symptoms in nfvPPA.............................. 187
  5.2.2 Inflexible predictions can explain auditory processing abnormalities in nfvPPA.............................. 189
  5.2.3 Inflexible predictions can explain receptive agrammatism in frontal aphasias.............................. 194

Summary............................................................................................................................................... 196

References............................................................................................................................................ 199

Appendix: Published work resulting from this thesis ....................................................................... 232
Table of Figures

Figure 1-1 MRI images for the amnestic Alzheimer’s disease vignette. ......................... 6
Figure 1-2 MRI images for the Posterior Cortical Atrophy vignette. .......................... 11
Figure 1-3 MRI images for the logopenic Primary Progressive Aphasia vignette........ 13
Figure 1-4 MRI images for the Semantic Dementia vignette.................................. 20
Figure 1-5 MRI images for the non-fluent variant Primary Progressive Aphasia vignette............................................................... 22
Figure 1-6 MRI images for the Dementia with Lewy Bodies vignette. ......................... 27
Figure 1-7 Serial structural MRI sections through the midbrain of a patient with PSP 30
Figure 1-8 MRI and SPECT images for the Corticobasal Syndrome vignette............ 32
Figure 1-9 DaT scan summary image for the Corticobasal Syndrome vignette. ....... 32
Figure 2-1 Movement parameters. ........................................................................... 52
Figure 2-2 The relationship between tau burden and age for each group................. 57
Figure 2-3 Connectivity matrices. ........................................................................... 58
Figure 2-4 Weighted degree. .................................................................................. 61
Figure 2-5 Alzheimer’s disease comparison of the three graph metrics representing the three principal hypotheses of hub vulnerability................................................. 63
Figure 2-6 PSP comparison of the three graph metrics representing the three principal hypotheses of hub vulnerability....................................................... 64
Figure 2-7 Unthresholded nodal connectivity strength. .......................................... 65
Figure 2-8 Weighted participation coefficient. ....................................................... 67
Figure 2-9 Clustering coefficient. ........................................................................... 69
Figure 2-10 Between-subjects analysis of the relationship between global tau burden and each graph metric at a network density of 6%, with the effect of age on tau burden partialled out................................................................. 73
Figure 2-11 Betweenness centrality. ...................................................................... 76
Figure 2-12 Closeness centrality (1/path length from each node to all other nodes)... 78
Figure 2-13 Local efficiency.................................................................................. 80
Figure 2-14 Eigenvector centrality. ........................................................................ 82
Figure 3-1 An illustration of the experimental motivation......................................... 98
Figure 3-2 Auditory symptoms in nfvPPA........................................................... 115
Figure 3-3 Behavioural results .......................................................................... 117
Figure 3-4 Total evoked power. .................................................................118
Figure 3-5 The effect of clarity. .................................................................120
Figure 3-6 The evoked effect of cue congruency. ......................................121
Figure 3-7 Evoked source space analysis. ..................................................123
Figure 3-8 Replication of experiment 1 with additional, neutral cues. ..........126
Figure 3-9 Example Bayesian model fits for a single subject in a single condition. ....129
Figure 3-10 Induced analysis after the written but before spoken word. ............131
Figure 3-11 Induced analysis after the spoken word. ....................................132
Figure 3-12 Scalp distribution and source reconstructions of induced effects. ....134
Figure 3-13 Coherence and connectivity analysis ........................................136
Figure 4-1 Group characteristics ...............................................................152
Figure 4-2 A plot for nfvPPA of difficulty ‘Understanding speech in a quiet room’ against ‘Grammatical form’ from the BDAE ........................................153
Figure 4-3 Composite pure tone audiograms for each group. .......................155
Figure 4-4 Artificial grammar structure and stimuli ......................................156
Figure 4-5 Distribution of decision variables. .............................................163
Figure 4-6 Group performance on sequence identification .............................166
Figure 4-7 Bias metrics by group for each language type .............................169
Figure 4-8 Correlation and clustering analysis .........................................170
Figure 5-1 Systems identification framework .............................................186
Figure 5-2 Basic auditory processing in nfvPPA, after Grube et al. (2016) ........190
Figure 5-3 Overall average adaptive tracks by group ..................................191
Figure 5-4 Individual nfvPPA performance profiles ....................................192
Table of Tables

Table 2-1 Participant demographics summary .......................................................... 49
Table 2-2 Detailed neuropsychological profiles of individual study participants. .... 50
Table 2-3 Sign tests ....................................................................................................... 83
Table 3-1 Demographic details of the experimental groups................................. 101
Table 3-2 Voxel-based morphometry regions of significant atrophy in nfvPPA...... 114
Table 3-3 Repeated measures ANOVA of evoked source power in frontal and temporal voxels of interest over the whole post auditory epoch................................. 124
Table 3-4 Repeated Measures ANOVA of behavioural data from experiment 1..... 125
Table 3-5 Repeated measures ANOVA of behavioural data from experiment 2...... 127
Table 4-1 Participant demographics........................................................................ 150
Table 4-2 Exposure and testing sequences. ................................................................. 157
Table 4-3 Single subject lesion percentages by region for the stroke group........... 161
Table 4-4 Statistical tests corresponding to Figure 4-6 and Figure 4-7. ................. 167
Table 4-5 Supplementary analysis of d’ discriminability.......................................... 181
Table 4-6 Supplementary analysis of d’ learning...................................................... 181
General Introduction

I undertook the work that comprises this thesis with three general aims in mind. The first was to demonstrate that systems neuroscience provides an effective and necessary translational bridge between the molecular biology of dementia and clinical trials of therapies and medications. The second was to demonstrate that studies of patients with neurodegenerative disease can further our understanding of normative brain function. In other words, I wanted to show that there is still a role for the ‘lesion study’ in experimental psychology. In chapters 3-5, I show that the application of modern imaging techniques and carefully designed psychophysics in patients with relatively selective but incomplete damage to part of the language processing network provides stronger evidence for causal interactions than associative studies in healthy individuals.

The third aim was that this thesis should represent a formative training in a comprehensive suite of computational, neuroimaging and psychophysical techniques. To this end, it intentionally comprises a diverse set of studies and investigations that employ a wide range of methods. Therefore, as well as making a meaningful contribution to scientific knowledge, this research I describe here is the philosophical foundation upon which I intend to build towards an independent research programme in the short timescales afforded by integrated clinical academic career pathways.

Although different in methods and approach, the chapters of this thesis provide complementary insights into the neurophysiology of dementia.

I begin with an introductory, clinically-focussed review (Chapter 1) that sets out the features, aetiology, management, epidemiology and prognosis of the dementias. This sets the scene for the experimental chapters of my thesis by placing the model diseases on which I place special emphasis within the context of the broader clinical challenge posed by dementia.

In the first experimental chapter (Chapter 2) I combine graph theoretical analyses of resting state functional MR imaging and PET imaging using the ligand AV-1451. This work followed from a number of ligand validation studies on which I was a co-author but that were not primarily my own work, and are therefore not reported here, which
together demonstrate AV-1451 to be effective at recapitulating the distribution of molecular pathology in vivo (Bevan-Jones et al., 2016, Bevan-Jones et al., 2017a, Passamonti et al., 2017). For this chapter I evaluated Alzheimer’s disease (AD) and progressive supranuclear palsy (PSP) as they represent prototypical neurodegenerative tauopathies with predominantly cortical and subcortical disease burdens respectively. I begin by considering how the connectivity of the brain might predispose to the contrasting patterns of neurodegeneration seen in these disorders, and empirically assess the hypotheses that have been proposed by cellular biologists and computational neuroscientists. Then, by taking a cross sectional approach, I go on to demonstrate those brain connectivity changes that occur with increasing tau burden. In this way, I am able to dissociate the causes of tau accumulation from its consequences.

The remainder of my thesis relies on the in-depth study of a disease called non-fluent variant Primary Progressive Aphasia (nfvPPA). This disease has a clear clinical phenotype of speech apraxia and agrammatism, associated with a focal pattern of mild atrophy in frontal lobes. The combination of relatively preserved cortical volume and well-preserved general cognition in early disease makes this an ideal model in which to study the neurophysiological consequences of neurodegeneration for connectivity and cognition.

The second, and largest, experimental chapter (Chapter 3) comprises an experiment in which patients with nfvPPA and matched controls performed a receptive language task while having their brain activity recorded with magnetoencephalography. I manipulated expectations and sensory detail to assess the role of top-down frontal contributions in predictive processes in speech perception. I am able to account for the behavioural psychophysics in a Bayesian perceptual inference framework, and to directly relate the outputs of these simulations to measured neural data. By relating evoked, induced and connectivity data, I demonstrate distant neural effects of the degeneration of frontal lobes that are not simply related to disconnection, and their associated behavioural consequences.

The final experimental chapter (Chapter 4) uses behavioural psychophysics to precisely define the ability of patients with nfvPPA to learn a novel, mixed-complexity artificial grammar designed to assess processing of increasingly complex sequencing relationships. I compare these abilities to patients with non-fluent aphasia due to stroke.
and neurological controls. In this way I am able to directly evaluate the impact of sequence modality and complexity on learning in agrammatic aphasics of two different aetiologies.

I conclude with a clinical synthesis chapter (Chapter 5), in which I explain how abnormal frontal lobe function and the consequential inflexible perceptual predictions can account for the subjective speech comprehension difficulties, auditory processing abnormalities and agrammatism of nfvPPA.

Finally, I summarise the way in which the thesis as a whole provides a translational bridge; relating cellular biology to molecular chemistry, molecular chemistry to neural circuits, neural circuits to core cognitive systems, and core cognitive systems to complex behaviours.
Chapter 1: The current clinical context of dementia

1.1 Preface

The aim of this narrative and introductory chapter is to present the current clinical context of dementia and neurodegenerative aphasia. I do this by describing the primary neurodegenerative conditions that impair cognition in terms of their clinical features, aetiology, management, epidemiology and prognosis. This sets the scene for the experimental chapters of my thesis by outlining how the diseases on which I place special emphasis relate to the broader clinical challenge posed by the dementias.

An extended version of this chapter, with additional material covering non-neurodegenerative memory disorders, has been accepted for publication as Chapter 25 of the upcoming Oxford Textbook of Neuropsychiatry, entitled Memory disorders and Dementias. The published version includes contributions from Dr Jeremy Isaacs of St George’s Hospital, London and Prof Michael Kopelman of King’s College London, but what I present here precedes their input and is entirely my own work. All of the figures in this chapter represent the clinical neuroimaging of patients whose care I have managed while conducting my PhD (with their consent).
1.2 Introduction

Dementia is a persistent disorder of the mental processes affecting more than one cognitive domain, which is generally but not universally progressive. It must be of sufficient severity to interfere with an individual’s adaptive function, and represent a decline from a previous level of function (McKhann et al., 2011). The clinical dementias can be classified either by their cognitive phenotype or by their aetiological origin, and there is not necessarily a linear mapping between the two.

This chapter is concerned with those disorders of mind or brain that result either in a primary dementia or a disorder of declarative memory. There are a large number of other conditions that result in secondary cognitive impairment or impair sensory or working memory that will not be discussed in detail. These include vascular insults and insufficiency, prion diseases, functional illness, traumatic brain injury, infections of the central nervous system, tumours, inflammatory and autoimmune disorders, demyelinating disorders, mitochondrial disorders, alcohol related disorders, side effects of drugs and toxins, obstructive sleep apnoea, nutritional deficiencies, and delirium.

The assessment of dementia is a multi-disciplinary pursuit, with clinical and diagnostic assessments undertaken by psychiatrists, psychologists, neurologists, geriatricians and, increasingly, specialist nurses, occupational therapists and other allied health professionals. A detailed history of the cognitive impairment is key, and is complemented by appropriate neurological examination and focussed investigation. For illustrative purposes the structural investigations presented will usually be magnetic resonance images (MRI), which are generally clinically preferred in the UK (NICE and SCIE, 2006 (updated March 2016)), but in many cases computed tomography (CT) would be clinically sufficient as the primary purpose of the scan is usually to exclude structural causes of the clinical syndrome (i.e. tumours, bleeds, infarcts and abscesses) (Cole, 1978). Similarly, case vignettes will include screening neuropsychology with the Addenbrooke’s Cognitive Examination third edition (ACE-III) (Hsieh et al., 2013) due to its simplicity of administration and favourable sensitivity and specificity (Larner,
2007, Larner, 2015), but similar clinical tools are available that might be more appropriate in particular practice settings.

**1.3 Core Progressive Dementias**

**1.3.1 Alzheimer’s Disease – Amnestic**

**Vignette**

A 70 year old gentleman presents to clinic, accompanied by his wife and daughter. His relatives tell you that over the past two years he has been becoming increasingly forgetful, especially for recent events. He has missed a number of appointments and recently became lost while driving an unfamiliar route. The gentleman himself feels that his relatives are over-emphasising his difficulties, but does admit to his memory not being as good as it once was. He is still able to manage all of his activities of daily living, but was unable to recall what he had for dinner yesterday. There are no abnormalities detected on physical or neurological examination. He scores 72 on an ACE-III (attention 12/18, memory 14/26, fluency 10/14, language 24/26, visuospatial 12/16). His MRI scan demonstrated mild global atrophy with disproportionate and symmetrical atrophy of the hippocampi relative to the rest of cortex (Figure 1-1)

![Figure 1-1 MRI images for the amnestic Alzheimer’s disease vignette.](image)

*Radiological convention; T1 weighted axial, coronal and sagittal projections. Bilateral, disproportionate atrophy of the hippocampi can be seen (arrows).*
Clinical Features

A number of diagnostic criteria are available for the diagnosis of Alzheimer’s disease, but the most commonly used is that jointly proposed by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (now known as the Alzheimer's Association) in 1984 (McKhann et al., 1984). These have stood the test of time, but there have been attempts to update them with more modern imaging techniques for both research (Dubois et al., 2007) and clinical (McKhann et al., 2011) purposes. The condition is characterised by an amnestic syndrome with insidious onset over months or years, without the core features of another of the conditions discussed later in this chapter.

Aetiology

In its ‘typical’ form, Alzheimer’s disease is a progressive disorder of episodic memory, reflecting the typical early involvement of the hippocampal formation. The first microscopically observable abnormalities are tau lesions (non-argyrophilic pretangle material, argyrophilic neuropil threads, neurofibrillary tangles), and later by amyloid-β deposition (Braak and Braak, 1995, Braak and Del Tredici, 2015). Accumulating evidence suggests that this pathological process begins decades before clinical onset (Braak and Del Tredici, 2015), but progresses extremely slowly; usually only crossing a threshold for clinically detectable abnormality many decades after it began. As the disease progresses, the accumulation of pathology accelerates and spreads to involve other cortical areas and therefore cognitive domains. Relatively rarely, the predominant burden of tau and amyloid pathology is outside of the hippocampal formation, resulting in the rarer presentations of Alzheimer’s disease discussed below.

Alzheimer’s disease is predominantly a sporadic disease, but risk and age of onset are influenced by a number of genetic polymorphisms (Harold et al., 2009). The strongest association is with the apolipoprotein E allele epsilon 4 (Saunders et al., 1993), the presence of which leads to a younger age of onset (Meyer et al., 1998). It is unclear whether this risk is primarily mediated by congenital alterations in brain activity and connectivity (Filippini et al., 2009) or a direct effect on beta amyloid (Bales et al., 1997). Around 2% of cases result from autosomal dominant mutations in genes coding for either the amyloid precursor protein (APP), or the gamma-secretase complex...
components presenilin 1 (PSEN1) and two (PSEN2) (Brouwers et al., 2008). Duplication of APP due to trisomy 21, and consequent overexpression of amyloid precursor protein, is the cause of early onset Alzheimer’s disease in Down’s syndrome (Rumble et al., 1989).

Management

At the time of writing, no effective disease modifying therapies are available for Alzheimer’s disease. By 2013, more than 200 drug candidates had been evaluated unsuccessfully (Becker and Greig, 2013). It is not clear whether these drugs were all ineffective because their mechanism of action was a flawed therapeutic strategy, or whether other factors were at play. Certainly, by the time disease is clinically evident underlying neuropathological abnormalities are advanced and progressing rapidly. It has been proposed that one or more of these candidates might still prove efficacious in pre-symptomatic disease, but clinical trials in this group are difficult because high-risk case selection is complex and measurement of progression in the absence of validated biomarkers requires long and expensive clinical trials (Fitzgerald, 2014). As well as drugs that specifically target Alzheimer-type pathology, novel agents are in development that target more general cellular stress response pathways to prevent toxic protein accumulation (Moreno et al., 2013, Halliday and Mallucci, 2014).

Therefore, at present, the aim of treatment is to manage symptoms. The mainstay of management for all of the diseases discussed in this chapter is to provide information and guidance to patients and their family about support available and the likely course of the disease to enable future planning. Assessment and information exchange should be holistic, and as a minimum should cover: advanced planning of decision making arrangements for health and finances (lasting power of attorney); financial matters and available benefits; housing and social arrangements and future plans; cognitive stimulation; carer support; and opportunities for pharmacological intervention.

Initial symptomatic treatment in mild to moderate disease is usually with a cholinesterase inhibitor (and, if possible, the cessation of anti-cholinergic medications (Lu and Tune, 2003)). On average these result in small improvements in cognitive test scores and measures of activities of daily living and behaviour (Birks, 2006). The three drugs in this class (donepezil, rivastigmine and galantamine) have similar efficacy
(Bullock et al., 2005). Which to initiate is therefore a clinical judgment based on practical concerns, side effect profile and cost. Donepezil has once daily dosing; while Rivastigmine has a patch preparation that can be useful in late disease if swallowing is impaired (this can be particularly useful in the Lewy body disorders). Donepezil seems to have a more tolerable side effect profile than rivastigmine (Bullock et al., 2005), but there is no difference in serious adverse effects. Common side effects include nausea, vomiting and diarrhoea – these effects can be dose dependent and it is therefore common practice to begin at a low dose, and slowly titrate upwards to the highest tolerable level. Serious adverse events are very rare, but caution should be exercised in the presence of cardiac rhythm abnormalities and patients with peptic ulcer disease (Schachter and Davis, 1999). In the absence of side-effects, it would be usual to continue these medications lifelong because of the risk of cognitive relapse on their cessation (Rainer et al., 2001).

Where dementia is moderate to severe (or in more mild disease where cholinesterase inhibitors cannot be tolerated), the addition of the NMDA receptor antagonist memantine should be considered (Raina et al., 2008). This has a small but robustly demonstrable benefit to measures of cognition, activities of daily living, behaviour and overall performance status (Tariot et al., 2004).

Pharmacological intervention for other specific neuropsychiatric symptoms should be addressed with caution. For example, while there is sometimes a role for low doses of antipsychotics in managing difficult behaviours that might have significant negative sequelae (such as jeopardising an otherwise successful residential placement), the side effect profile of these drugs means that it is normally preferable to address these issues with social interventions and targeted carer support.

Low mood is a frequent complaint in dementia. Unfortunately there is converging evidence that commonly used antidepressant medications are broadly ineffective at treating low mood in Alzheimer’s disease, while carrying a side effect burden (Banerjee et al., 2011). Mirtazapine has been shown to improve carer quality of life, primarily by improving appetite and sleep, while high-dose citalopram was effective for agitation (Porsteinsson et al., 2014). The specific indication for treatment should therefore be carefully considered.
Epidemiology

Alzheimer’s disease is the most common dementia in the western world, representing approximately 70% of the 4.6 million global annual incident cases (Ferri et al., 2005). Prevalence is felt to be rising especially rapidly in the developing world, as life expectancy increases (Reitz et al., 2011). While age-for-age dementia incidence is falling in the UK (Matthews et al., 2013, Matthews et al., 2016), the ageing population means that dementia continues to rise in both prevalence and incidence. Prevalence increases rapidly with age, from <5% of those below age 75, to 10% at age 80-84, 20% at 85-89 and 30% over 90.

Prognosis

Alzheimer’s disease is a relentlessly progressive condition that alters behaviour (Reisberg et al., 1996), induces dependency and shortens life expectancy. The most powerful predictors of life expectancy from diagnosis are age, sex, degree of cognitive impairment at diagnosis, and magnitude of cognitive decline over the first year (Larson et al., 2004). Median life expectancy for 70 year old women with Alzheimer’s disease is 8 years, compared to 16 years for those without. For men at 70, it is 4.4 years compared to 12.4 years without. Diagnosis at an older age has less impact on survival – prognosis for Alzheimer’s sufferers at age 90 is 2-3 years, compared to 3-4 years without. Mini-mental state exam scores below 24/30 at diagnosis predict poorer prognosis, with those below 17 having a particularly high hazard ratio for death (2.6x that of an MMSE > 24). A decline of more than 5 points in the MMSE over the first year results in a hazard ratio of 1.6. Recent suggestions of multimodal prognostication using quantitative neuroimaging and CSF analysis have not yet found clinical utility, but hold promise for research applications (Perrin et al., 2009).

1.3.2 Alzheimer’s Disease – Posterior Cortical Atrophy – Benson’s Syndrome

Vignette

A 58 year old man presents to clinic with his wife. Two years ago he noticed difficulty with using money, being unable to easily distinguish between coins without feeling them. A year later he began to have problems with not noticing hazards while driving.
His wife describes him, “looking but not seeing.” He has been assessed by an optician and an ophthalmologist, both of whom said that his visual acuity was normal, but suggested that there might be a neurological cause for his problems. On examination you note simultagnosia, bilateral dyspraxia, dyscalculia, dysarithmetria and visual working memory difficulties. His speech is also slightly hesitant, and he displays some anomia with preserved semantic knowledge during confrontational naming. There are no abnormalities on neurological examination, and specifically there is no movement disorder. ACE-III is 82/100 (attention 18/18, memory 22/26, fluency 12/14, language 22/26, visuospatial 8/16). His MRI scan demonstrated global atrophy especially severe in the parietal lobes (Figure 1-2).

Figure 1-2 MRI images for the Posterior Cortical Atrophy vignette.
Radiological convention. Proton density weighted axial projection; T1 weighted coronal projections through frontal and then parietal regions; T1 weighted sagittal projection. Disproportionate, bilateral parietal lobe atrophy can be seen.

Clinical Features
Reflecting the biparietal predominance of pathological damage (Ossenkoppele et al., 2015), posterior cortical atrophy is a primary disorder of the dorsal ‘where’ visual stream (Ungerleider and Mishkin, 1982, McMonagle et al., 2006), leading to visuospatial disorientation with or without a broader biparietal syndrome of ideomotor and gesture copy apraxia, Gerstmann syndrome (dysgraphia, dyscalculia, finger agnosia and left-right confusion) and Bálint’s syndrome (optic ataxia, ocular apraxia and simultagnosia). An associated occipitotemporal syndrome of alexia and apperceptive agnosia (including prosopagnosia) is also common, as is the later emergence of a parietal, Alzheimer-type language deficit (see next section on logopenic and mixed primary progressive aphasia).

Prognosis
Although there is a clinical perception that the prognosis of posterior cortical atrophy is better than that of amnestic Alzheimer’s disease, this seems to be a result of the generally younger age of presentation and the often shorter duration of symptoms before coming to medical attention. When age matched groups are studied, there is no significant difference between typical and focal Alzheimer’s disease in time from first symptoms to death (mean 9.7 years) (Alladi et al., 2007).

Epidemiology and Management

There have been no robust epidemiological or prevalence studies of focal presentations of Alzheimer’s disease. The clinical perception is that even if all of the focal subtypes are combined, they are still far less common than typical Alzheimer’s disease. Focal presentations tend to present at an earlier age than typical AD, but it is not clear whether this represents a referral or case selection bias. Medical management is generally similar to typical Alzheimer’s disease, as the symptoms become more phenotypically similar as disease progresses. The additional consideration in posterior cortical atrophy is management of the visual symptoms, which can be very disabling. Driving should cease at an early stage. Some patients have found it beneficial to become registered as legally blind, as in the UK this facilitates access to a number of benefits and concessions, as well as practical support from charitable bodies. This application can require careful wording, as patients with cortical blindness of this type will have good visual acuity and the visual field might appear normal to single-target confrontation. An ophthalmologist’s report will be required, and the criterion that can usually be applied is that which states: “If you have a good visual acuity, you will usually have had to have lost a large part of your visual field to be certified as severely sight impaired (blind) or sight impaired (partially sighted).” It should be stressed that, while the detection of single targets can be in-tact, analysis of the visual field as a whole is catastrophically impaired.

1.3.3 Alzheimer’s Disease – Logopenic and Mixed Primary Progressive Aphasia

Vignette

A 59 year old right-handed gentleman presents to clinic with an 18 month history of subjective speech disturbance. He describes difficulty in finding words, such that he
becomes derailed half way through a sentence. He is clear that he can bring to mind the object or concept and give a description of it, but feels as if the word itself is stuck ‘at the tip of his tongue’. He has given up work as a local government planning officer because of this difficulty, although his managers had not felt that his performance had deteriorated. Conversationally he is fluent and his speech is full of content, but there are pauses as he gropes for words, and sometimes the train of thought is lost. There are some problems with confrontational naming, but he is able to give a semantically detailed description of these items. He finds it difficult to repeat long sentences, giving close approximations that carry the gist and are grammatically correct. He scores 89/100 on the ACE-III (Attention 18/18, Memory 19/26, Fluency 10/14, Language 26/26, Visuospatial 16/16); note that the language testing on ACE-III is relatively simple, with no long sentence repetitions, and so the symptomatic deficits in this domain were not evident on bedside cognitive testing. His MRI demonstrates mild global atrophy, with a parietal predominance (Figure 1-3).

Figure 1-3 MRI images for the logopenic Primary Progressive Aphasia vignette. Radiological convention. T1 weighted coronal and sagittal projections. Subtle atrophy of hippocampi and parietal lobe can be seen, with a left sided predominance.

Clinical Features

Aphasia is very common in typical Alzheimer’s disease (Cummings et al., 1985), but in rare cases it can be the leading or only clinical feature. The core clinical feature of this condition is a deficit in responsive and confrontational naming with preserved semantic knowledge. This can often be subtle at presentation, as patients notice the problem early, frustrated by pauses in their speech as they find themselves groping for words.
The diagnostic criteria for logopenic variant primary progressive aphasia are the presence of impaired single-word retrieval in both spontaneous speech and naming together with impaired repetition of sentences and phrases in a length dependent manner, as well as three out of the following: phonological errors, spared single word comprehension and object knowledge, spared motor speech, and absence of agrammatism (Gorno-Tempini et al., 2011). These features are strongly supportive of Alzheimer’s pathology involving left parietal lobe (Gorno-Tempini et al., 2004). In recent years, it has been suggested that these diagnostic criteria are too narrow. There is converging evidence that any primary progressive aphasia that does not meet diagnostic criteria for semantic dementia or progressive non-fluent aphasia (i.e. a mixed aphasia), is also likely to represent Alzheimer’s pathology in left parietal lobe (Sajjadi et al., 2012, Sajjadi et al., 2014), especially if confrontational anomia or amnesia is present (Alladi et al., 2007).

1.3.4 Alzheimer’s Disease – Dysexecutive/Behavioural Variant

Clinical Features

It can be very difficult to clinically distinguish between a behavioural variant of Alzheimer’s disease and the behavioural variant of frontotemporal dementia (bvFTD). Presentations can be very similar (see vignette and clinical features of bvFTD below) (Woodward et al., 2010). The clinical syndrome reflects the distribution of neuronal loss rather than its aetiology, with distinct regional atrophy patterns found in syndromes characterised by disinhibition, apathy or aberrant motor behaviours (Rosen et al., 2005). The presence of significantly impaired episodic memory is not sufficient evidence that one is dealing with a dementia of the Alzheimer’s type, as there is now converging evidence for memory involvement in a significant proportion of bvFTD (Hornberger et al., 2010, Hornberger and Piguet, 2012), but should prompt the clinician to consider further investigation with detailed neuropsychological, neuroimaging and/or neuropathological assessment (Larner, 2006). While an educated guess as to the likely aetiology can be made, it should be borne in mind that without pathological or biomarker confirmation at least 5-10% of seemingly typical cases are misclassified (Graham et al., 2005, Alladi et al., 2007, Rascovsky et al., 2011).
Management

While cholinesterase inhibitors have some efficacy in managing the neuropsychiatric manifestations of Alzheimer’s disease (Wynn and Cummings, 2004), they have been reported to worsen behavioural symptoms in frontotemporal dementia (Mendez et al., 2007, Kimura and Takamatsu, 2013). Similarly, while Memantine had promising initial anecdotal reports of efficacy in bvFTD (Swanberg, 2007), two randomised controlled clinical trials proved negative (Vercelletto et al., 2011, Boxer et al., 2013). Given the practical difficulties in distinguishing these two syndromes, circumspection and careful assessment of response should be employed in prescribing. Strategies for the management of bvFTD are discussed in the next section.

1.3.5 Frontotemporal Dementia – Behavioural Variant – bvFTD

Vignette

A 77 year old right handed lady is referred to clinic by her family. They report that she has become progressively more ‘childish’ over the past two years. She is inappropriately familiar with strangers, and displays total disregard for etiquette such as table manners. You notice that her clothes are dirty and that she is wearing three scarves, two pairs of trousers and two bras. Her husband, who is well kempt, reports that she insists on re-wearing dirty clothes and that he is unable to intervene. The patient tells you that there is a man living in her attic; she has never seen him, but knows that he is there. On examination you note the presence of primitive reflexes (grasp, palmomental, pout), as well as deltoid fasciculations, thinning of the first dorsal interosseous and thenar eminence bilaterally, slow tongue movements and hyperreflexia. ACE-III is 85/100 (Attention 16/18, Memory 21/26, Fluency 10/14, Language 21/26, Visuospatial 16/16). Her language deficits are in naming, and you note that she has some mild semantic impairment for these items.

Clinical Features

The most recent diagnostic criteria allow for the diagnosis of ‘possible’ bvFTD on clinical grounds in the presence of three of the following six features: disinhibition, apathy, loss of empathy, perseverative behaviours, hyperorality (an excessively sweet
tooth) and dysexecutive neuropsychological profile (Rascovsky et al., 2011). The level of certainty is increased to ‘probable’ in the presence of functional disability and characteristic neuroimaging, while ‘definite’ diagnosis requires either histopathological confirmation or a pathogenic mutation. It is especially important to assess for progression of symptoms and neuroimaging changes, as a significant proportion of cases diagnosed on clinical features alone fail to progress in terms of clinical syndrome or atrophy (Hornberger et al., 2009, Kipps et al., 2010). It is unclear whether these cases represent atypical, late-onset neuropsychiatric disease (Belin et al., 2015) or are simply retiring males spending more time with, often new, life partners. When progression is established (Mioshi and Hodges, 2009), these criteria have a sensitivity of 72-95% and specificity of 82-95% (Harris et al., 2013), with the majority of misclassifications being with frontal Alzheimer-type pathology as discussed above.

Over time, the evolution of the syndrome to include semantic deficits is common, reflecting a shared aetiology with semantic dementia. A proportion of patients with bvFTD will go on to develop progressive supranuclear palsy (Kaat et al., 2007) or, as here, motor neurone disease.

Aetiology

The clinical syndrome of FTD represents several histopathological patterns of frontotemporal lobar degeneration. 40%-50% of cases display inclusions of the microtubule-binding protein tau (Mackenzie et al., 2010) with either three or four microtubule-binding repeats (Cairns et al., 2007). Approximately 40-50% of cases have inclusions of TDP-43 (43 kDa TAR DNA-binding protein) (Piguet et al., 2011) and 5% inclusions of FUS (fused in sarcoma) (Seelaar et al., 2010a). These proteins are both involved in RNA-processing, but their pathogenic mechanisms are not yet known. The remaining 5% of cases with typical phenotype and distribution of neuronal loss (and without Alzheimer-type or Lewy body pathology) are negative for tau, TDP-43 and FUS, suggesting an as-yet unknown pathogenic mechanism (Mackenzie et al., 2008).

bvFTD is highly heritable, with up to 50% of patients having a family history of the disorder (a far higher proportion than the other FTD syndromes, which do not generally run in families) (Rosso et al., 2003). A single causative gene can be
demonstrated in approximately half of these cases, with the proportion rising in clear autosomal dominant inheritance patterns with multiple family members affected (Rohrer et al., 2009). It is now possible to test for most of the common exon mutations on a single chip; MAPT and GRN each account for 5-11% of FTD, while TARDBP, FUS, CHMP2B, TBK1, VCP, SQSTM1 and hnRMP1a account for progressively fewer cases and tend to be associated with additional clinical features (Seelaar et al., 2010b). At least as common as any of these is a hexanucleotide repeat expansion in the intron C9ORF72. This must be separately tested, and is particularly suggested by the presence of motor neurone disease (DeJesus-Hernandez et al., 2011, Renton et al., 2011) or prominent psychiatric features in the proband or family members (Arighi et al., 2012, Snowden et al., 2012). While C9ORF72 has autosomal dominant inheritance, it is incompletely penetrant and manifests differently within a family, as motor neurone disease, psychiatric disorders and bvFTD alone or in combination. In the vignette, the combination of bvFTD, motor neurone disease and delusions would strongly suggest C9ORF72, even in the absence of a family history.

Management

In contrast to AD, there is no role for cholinesterase inhibitors in bvFTD (Mendez et al., 2007, Kimura and Takamatsu, 2013). Similarly, while Memantine had promising initial anecdotal reports of efficacy (Swanberg, 2007), two randomised controlled clinical trials proved negative (Vercelletto et al., 2011, Boxer et al., 2013).

One medication with evidence for efficacy in relieving the behavioural symptoms of bvFTD is trazodone (Lebert et al., 2004), from which sustained benefit can often be obtained (Lebert, 2006). These benefits are not always observed, and the decision to continue with this medication should be made by the prescriber based on a comprehensive assessment of both patient and carer experience. A possible disease-modifying mechanism for trazodone has been proposed through actions on the cellular unfolded protein response (Halliday et al., 2017), however any conclusions about clinical efficacy would be premature without carefully conducted clinical trials (Halliday and Mallucci, 2017).

The mainstay of treatment is non-pharmacological. Abnormal behaviours can be challenging, especially for members of the public. It is important to consider whether
any community stakeholders would benefit from being aware of the diagnosis, and suggesting this to patients and their families. It is often appropriate for families to inform local law enforcement services and frequently visited shops.

Driving can be a particularly thorny issue – patients with bvFTD, more than any other group in the dementia clinic, lack insight into their symptoms and the problems they cause in judgment. While it is usual practice in the UK to tell patients that they must inform the DVLA of their diagnosis, in bvFTD it is often necessary to firmly advise against driving and directly inform the DVLA of this advice. Family members can be a useful guide towards the necessity of taking such action.

Hyperorality and dietary changes are often a significant issue. Overeating can represent a simple utilisation behaviour, in which case it can be fairly easily addressed by selective purchasing and locking cupboards. In some individuals, however, obtaining sweet foods becomes an overarching goal-directed behaviour, and weight gain can be dramatic (Ikeda et al., 2002, Seeley et al., 2005). Social interventions are often insufficient in these cases, and pharmacological intervention required. Diabetes is often comorbid, and metformin can be a useful suppressor of appetite. Similarly if myoclonus, tremor or seizures are present, topiramate can be considered. Limited or repetitive diets can be similarly damaging, and in such cases potential vitamin deficiencies should be addressed lest a diet of biscuits and fizzy drinks lead to scurvy. In advanced disease, weight gain gives way to weight loss as more than 80% of patients become unable to eat independently (Diehl-Schmid et al., 2017).

**Epidemiology**

The behavioural variant of FTD has a significantly younger age of onset than the other disorders in the FTLD spectrum (Coyle-Gilchrist et al., 2016), with the majority of both genetic and sporadic cases presenting before the age of 65 (peak incidence across FTLD as a whole is between age 65 and 80). Incidence of FTLD in the UK is about 10/100,000 person-years, with bvFTD representing approximately a quarter of these cases (SD and nfvPPA together represent another quarter, with the other half being CBS or PSP). When adjusted for higher female life expectancy, there is no significant gender bias in bvFTD or in FTLD as a whole.

**Prognosis**
Mean survival from symptom onset is approximately 8 years (in the absence of signs of motor-neurone disease). The clinical syndrome evolves significantly through this period, as the over-activity, social inappropriateness and hypersexuality that can be so distressing in early disease frequently give way to a syndrome of apathy and semantic impairment after the first few years. This should be recognised by dynamic symptom management, with the cessation of no-longer-necessary medications and pro-active interventions given equal priority.

1.3.6 Frontotemporal Dementia – Semantic Dementia - SD

Vignette

A 70 year old right-handed man is brought to clinic by his wife. He has not noticed any problems, and is unsure why he has been brought to clinic. His wife says that she first noticed something wrong a year ago, when they were perusing a restaurant menu and her husband asked her what asparagus was. Since then, his understanding of words and concepts, both written and spoken, has gradually declined. She relates that some activities are completely unchanged, for example her husband is still able to drive and route find, and still enjoys playing bridge to a relatively high standard. The patient’s conversational speech is fluent and grammatical but empty of content words. He is unable to name lower frequency animals, and nor is he able to tell you which of several model animals is dangerous or edible. When asked to draw animals he produces images that lack distinctive features – for example a camel without a hump and an elephant without a trunk (Bozeat et al., 2003). Similarly, he cannot name or demonstrate the use of some kitchen implements. He is able to repeat long words and sentences without difficulty, but has poor understanding. When reading aloud he pronounces irregular words phonetically (surface dyslexia) (Woollams et al., 2007). ACE-III is 70/100 (attention 18/18, memory 22/26, fluency 2/14, language 12/26, visuospatial 16/16). His MRI scan demonstrates severe, selective atrophy of the left temporal lobe (Figure 1-4).
Clinical features

SD is a stereotyped syndrome of progressive loss of semantic knowledge associated with left sided temporal lobe atrophy (Warrington, 1975, Basso et al., 1988, Snowden et al., 1989, Hodges et al., 1992). It is often classified together with the other neurodegenerative aphasias (nfvPPA and logopenic Alzheimer’s disease), as fluent, content-less speech is a leading symptom; in this context it is named the semantic variant of primary progressive aphasia, svPPA. The aphasia is, however, a consequence of a broader loss of semantic knowledge for objects. This can be best conceptualised by an example: if a patient with logopenic Alzheimer’s disease is presented with a model lion, they might not be able to tell you that it is a lion (displaying anomia), but they would be able to give semantic detail about lions, for example that they are dangerous and live in Africa. A patient with SD might similarly lack the word ‘lion’, but they would also lack all knowledge about lions. Therefore, although they do indeed have an agnosia for words, this amodal deficit extends to all semantic domains and is not, primarily, an aphasia (Bozeat et al., 2000, Fairhall and Caramazza, 2013).

In fact, it is possible that a patient with SD will tell you that what they are looking at is a cat, and that they have one at home. This reflects the postulated mechanism for semantic memory, whereby knowledge about objects is stored as a commonly observed ‘template’, with specific differentiating flags. For example, an elephant might be stored
as a commonly observed large quadruped such as an horse, with the additional flags ‘has a trunk’, ‘lives in Africa or Asia’, ‘has big ears’, ‘is very large’ etc. (Snowden et al., 2001). This can be clinically observed by asking patients with SD to draw animals with distinguishing features. The more unique the feature, the more likely it is to be omitted – for example a patient with SD might draw an elephant without a trunk, a camel without a hump, or a duck with four legs (Bozeat et al., 2003).

Behavioural features are common (Seeley et al., 2005, Coyle-Gilchrist et al., 2016, Lansdall et al., 2017), especially later in the disease course when the right temporal lobe becomes involved. These overlap with bvFTD, but are more likely to comprise compulsive, repetitive behaviours and disinhibition than apathy and indiscriminate eating, reflecting a temporal rather than frontal predominance of atrophy (Snowden et al., 2001). The syndrome of predominant right sided temporal atrophy is dominated by this behavioural phenotype (Chan et al., 2009), often in combination with prosopagnosia (loss of semantic knowledge of familiar people) (Evans et al., 1995). Similarly, as temporal lobe atrophy becomes bilateral, an additional SD phenotype often develops.

Overall, there can be striking preservation of ability outside of semantic and behavioural domains. It is not uncommon for a patient to demonstrate catastrophic loss of object knowledge, but retain complex procedural skills and navigational ability. Schemata can be well preserved, but sometimes break down with object substitution – for example proficiently making a cup of tea but then putting the milk in the cupboard and tea bags in the fridge.

**Aetiology**

Semantic dementia is almost always associated with TDP-43 related histopathology affecting the temporal lobes, with a left sided predominance. In contrast to bvFTD it does not show significant heritability, and indeed at least three patients seen in the Cambridge clinic are known to have had SD and an unaffected monozygotic twin (personal communication from Prof Karalyn Patterson).
1.3.7 Frontotemporal Dementia – non-fluent variant Primary Progressive Aphasia – nfvPPA/PNFA

**Vignette**

A 63 year old right-handed lady presents to clinic with a one year history of a progressively worsening difficulty in getting her words out. She finds this extremely frustrating. She is very clear that she knows what the word is that she wants to say, but she is unable to articulate it. She has not noticed any other problems. Her conversational speech is effortful, with phonological paraphasias being especially noticeable with words requiring articulatory agility. It is also telegraphic, lacking connecting words and some grammatical form. She performs poorly at parsing grammatically complex sentences, but demonstrates flawless confrontational naming and semantic knowledge. You ask for a sample of her writing, and note that this too contains grammatical errors. She scores 96/100 on the ACE-III (Attention 18/18, Memory 25/26, Fluency 12/14, Language 25/26, Visuospatial 16/16). Her MRI scan appeared normal (Figure 1-5).

![MRI images for the non-fluent variant Primary Progressive Aphasia vignette.](image)

_Radiological convention. T2 weighted axial projection; T1 weighted coronal and sagittal projections. Normal appearances for age can be seen, with good hippocampal bulk._

**Clinical Features**

Non-fluent variant Primary Progressive Aphasia (nfvPPA), formerly known as Progressive Non-Fluent Aphasia (PNFA), is an adult onset neurodegenerative aphasia...
characterised by apraxia of speech and/or agrammatism (Gorno-Tempini et al., 2011). Importantly, patients should not have word findings difficulties (anomia) or a general cognitive impairment in the early stages of disease; if this is present, a diagnosis of mixed primary progressive aphasia should be made and the underlying neuropathology is almost always of the Alzheimer’s type (Sajjadi et al., 2012, Sajjadi et al., 2014). Sentence repetition is impaired where articulatory agility is required or the grammatical structure is complex. Semantic deficits should not be present in early disease, but these and/or behavioural features can evolve later. Yes/no confusion is a common feature in relatively early disease, and digit span can be markedly shortened due to impairment of the phonological loop (Baddeley and Hitch, 1974). Neuroimaging changes can be very subtle in early disease, often appearing normal in the single subject, but predominant atrophy is in left frontal lobe (Rogalski et al., 2011, Grube et al., 2016).

Although nfvPPA is commonly thought of as an expressive aphasia, the neurodegenerative equivalent of Broca’s syndrome, patients also frequently complain that perceiving speech is effortful, even in optimal listening environments. Many have sought hearing aids, but been told that they do not have a peripheral deficit. It is thought that this reflects a double-hit to auditory processing, with impairments in basic auditory processing (Goll et al., 2010, Grube et al., 2016) compounded by damage to frontal regions making predictions to ‘fill-in’ unclear speech (see Chapter 5).

The diagnostic criteria require the presence of only one of apraxia of speech and agrammatism. This has led some authors to propose that nfvPPA should be clinically subdivided into an agrammatic variant (agPPA) and an apraxic variant (primary progressive apraxia of speech, PPAOS) (Josephs et al., 2012, Josephs et al., 2013). The justification for this division has been the observation that those presenting with PPAOS are more likely to develop a motor syndrome consistent with PSP or CBS later in their illness (Josephs et al., 2005) and have a neuropathological diagnosis of tau rather than TDP-43 pathology (Josephs et al., 2006). However, this subdivision has not gained widespread acceptance in the aphasia community, and the new research diagnostic criteria for PSP acknowledge that apraxic speech and agrammatism either individually or in combination can be associated with underlying PSP pathology (Höglinger et al., 2017). There are several reasons for this:
1) There is a clear spectrum to nfvPPA, with some patients displaying predominant apraxia of speech, and others having marked difficulties with both expressive and receptive grammar, but all patients displaying deficits in both domains on careful testing.

2) Fewer than half of patients with PPAOS will develop a movement disorder, and even if this does occur this transformation often emerges five or more years into the illness (Josephs et al., 2012, Josephs et al., 2014). Similarly, movement disorders sometimes occur later in the illness of those who present with agrammatism.

3) While there is a statistical trend for symptom type at onset to predict progression and pathology, even in the presence of a movement disorder, predominant TDP-43 pathology in the absence of tau has been observed. I have encountered one such case in my own practice, through donation to the Cambridge Brain Bank.

4) Although there are differences in patterns of atrophy between PPAOS and agPPA at the group level (Josephs et al., 2013), these are relatively subtle. Changes in cortical thickness in left inferior frontal gyrus correlate with grammatical processing, while those in inferior frontal sulcus correlate with fluency (Rogalski et al., 2011). These are neighbouring regions, and the differences are certainly less marked than those between PPAOS and PSP (Whitwell et al., 2013). Regardless of diagnostic subgroup, progression of aphasic symptoms is associated with neurodegeneration in the language network, while progression of motor symptoms is associated with neurodegeneration outside of the language network, in motor cortex and brainstem (Tetzloff et al., 2017).

5) nfvPPA is already a very rare syndrome, and further subdivision would make clinician education and meaningful research more difficult.
**Aetiology**

The syndrome of nfvPPA predicts atrophy centred on left frontal lobe, but not the nature of the causative pathology. The majority of cases represent tauopathies, but a significant minority have TDP-43 related disease (Kertesz et al., 2005, Josephs et al., 2006, Knibb et al., 2006a, Knibb et al., 2006b, Mesulam et al., 2014). Neuropathological series vary markedly (from 0-40%) in the proportion of nfvPPA found to have Alzheimer’s neuropathology. This likely represents differences in local diagnostic practice, perhaps in terms of the stringency with which anomic patients are excluded from nfvPPA diagnosis.

**Management and prognosis**

Patients with nfvPPA are intensely frustrated by their symptoms, and present early in their disease course (Coyle-Gilchrist et al., 2016). The cornerstone of therapy is explanation of symptoms and practical management of the aphasia; there is no role for medications unless additional motor or behavioural features emerge later in the disease. It is often fruitful to approach nfvPPA in the same way as Broca’s aphasia, with the provision of communication aids and explanation cards. Some patients even find it simpler to tell shopkeepers and acquaintances that they have had a stroke, as this encapsulates their problems more quickly than an explanation of an unusual neurodegeneration. Speech therapy can be helpful in teaching strategies to avoid words with challenging articulation or sentences with ambiguous grammar, but attempts at rehabilitation are rarely fruitful.

Yes/no confusion can be challenging and perplexing to carers, and should be explicitly acknowledged and explained. Strategies for dealing with this include asking patients to include a gesture as well as a verbal response (for example thumbs up and saying yes together so that discordance can be more easily recognised), and training carers to present non-binary choices.

Survival from symptom onset averages 9 years (similar to bvFTD), but as symptoms are troublesome presentation is often earlier. Age of onset is also older than bvFTD, with peak rates observed between age 65 and 80 (Coyle-Gilchrist et al., 2016). Even when patients present at a young age, a genetic cause is rarely found; while patients with GRN mutations often have non-fluent speech, they also have anomia and single
word comprehension problems not consistent with a diagnosis of nfvPPA (Rohrer et al., 2010).

1.4 Dementias associated with Movement Disorders

The neuropsychiatry of movement disorders could form a thesis in itself (see, for example, (Ghosh, 2011), leading to (Ghosh et al., 2012)). The discussion here will therefore be restricted to those clinical features that aid the differential diagnosis of these disorders in the cognitive clinic.

1.4.1 Lewy Body related – Dementia with Lewy Bodies – DLB

Vignette

A 73 year old lady is referred by her family doctor because of a three year history of hallucinations. These occur daily, primarily when she is tired or the lighting is poor. The most common manifestation is that she will make three or four coffees for people sitting on the sofa, and her husband will have to remind her that it is only the two of them in the house. She also frequently hallucinates individuals in the garden, and on one occasion thought that they were stealing her furniture. This caused some minor distress. She has good retrospective insight into the fact that these people are not real, but finds it very confusing at the time. If her husband reassures her that they are just hallucinations, she is able to understand this. Her cognition is fluctuant, with good days and bad days. She has REM sleep behaviour disorder. On examination you elicit a mild Parkinsonian motor syndrome, which is not symptomatic, and her sense of smell is poor. Despite only scoring 53/100 on an ACE-III (attention 11/18, memory 12/26, fluency 7/14, language 14/26, visuospatial 9/16), she is still able to travel to and from town alone on the bus once a week, where she follows a fixed routine of going to Marks and Spencer’s and coming home again. Her MRI scan demonstrated mild global atrophy (Figure 1-6).
Clinical features

DLB is a dementia characterised by fluctuating cognition, attention and alertness and/or recurrent visual hallucinations. The diagnosis is upgraded from ‘possible’ to ‘probable’ by the presence of REM sleep behaviour disorder, severe neuroleptic sensitivity or abnormal dopamine transporter activity on nuclear imaging (McKeith et al., 2005). Impairments of visuospatial, executive and attentional function can be particularly prominent. Other supportive features include repeated falls and syncope, transient unexplained losses of consciousness, autonomic dysfunction, delusions, depression, and the absence of hippocampal atrophy on structural imaging. DLB is diagnosed when this characteristic cognitive syndrome develops either in the absence of, or within one year of the presentation of a movement disorder typical of idiopathic Parkinson’s disease. The evolution of a movement disorder is not inevitable, as a proportion of patients dying with DLB have relative preservation of nigral dopaminergic neurones and consequently normal dopamine transporter imaging (Colloby et al., 2012). It is hypothesised, but not demonstrated, that this group might show less severe adverse neuroleptic reactions.

1.4.2 Lewy Body related – Parkinson’s Disease Dementia - PDD

Parkinson’s disease dementia is diagnosed when the syndrome of dementia with Lewy Bodies described above evolves more than one year into typical idiopathic Parkinson’s disease (McKeith et al., 2005). There is evolving consensus that DLB and PDD represent the same underlying disease, with the phenotypic presentation reflecting a
spectrum of relative predominance of Lewy body pathology in cortex and brainstem (McKeith and Burn, 2000, McKeith, 2009). Patients with Parkinson’s disease have quadruple the relative risk of incident dementia of age matched controls (Hobson and Meara, 2004). Although severity and duration of Parkinsonian symptoms are risk factors in development of dementia, the strongest risk factor is the patient’s current age, with these factors having a multiplicative effect (Levy et al., 2002). Not all dementia developing in Parkinson’s disease is PDD – Alzheimer’s pathology is found in up to two thirds of patients with Lewy body disease (Barker et al., 2002), and it is therefore perfectly reasonable to make concomitant diagnoses of idiopathic Parkinson’s disease and either Alzheimer’s disease or mixed dementia in the presence of a characteristic amnestic syndrome.

1.4.3 Progressive Supranuclear Palsy - PSP

Vignette

A 58 year old man is referred because he has begun to struggle at work. He is a maintenance engineer, and over the past year has been purchasing an increasing number of unnecessary instruments and gadgets. More recently he has made some inappropriate sexual comments and, when challenged, he has become aggressive. In clinic you note that he is inappropriately familiar and labile in his mood, quickly alternating between jocularity and confrontation. His speech is adynamic, with palillogia at times (repetition of a single word with no words in between). The conversation is slow. He has some insight into his changed personality, describing himself as previously being a reserved individual but acknowledging that he now says whatever is on his mind even if it hurts others. He also admits to a sweet tooth, stating that he is eating more cakes and would eat a whole packet given the chance. On neurological examination the only abnormality is slowed saccades, especially in the vertical plane, but with no restriction in the extent of eye movements. He scores 70/100 on ACE-III, with predominant deficits in attention and verbal fluency (attention 9/18, memory 20/26, fluency 3/14, language 24/26, visuospatial 14/16). His MRI scan demonstrated mild global atrophy, with disproportionate volume loss in the midbrain (see Figure 1-7 for longitudinal imaging from a different patient). You follow him up, and by one year he has begun to suffer frequent unprovoked falls characterised by
postural instability and an inability to correct in time. A supranuclear gaze palsy has now emerged, with almost complete restriction of voluntary eye movements in the vertical plane. By two years, he is confined to a wheelchair and beginning to have significant difficulties with chewing and swallowing. He has severe axial rigidity, but tone in his limbs remains relatively normal.

**Clinical Features**

PSP is a highly protean disease, with only around a quarter of pathologically confirmed cases presenting initially with the classic Richardson’s syndrome of falls, postural instability, cognitive dysfunction and abnormal eye movements (Respondek et al., 2014). Previous diagnostic criteria (Litvan et al., 1996) did not reflect this, and consequently suffered from limited sensitivity (Hughes et al., 2002, Osaki et al., 2004) at first presentation. To address this, the Movement Disorder Society have recently published new diagnostic criteria recognising early forms of the disease (Höglinger et al., 2017). These criteria split ‘variant’ PSP into a new classification based on predominant symptomatology, recognising onsets with movement disorders, cerebellar disorders, primary lateral sclerosis, behavioural features and abnormal speech (Respondek and Höglinger, 2016). As illustrated in the vignette, it is important to remember that up to 20% of PSP initially presents with a frontal cognitive syndrome that is sometimes indistinguishable from bvFTD until the evolution of the characteristic movement disorder (Kaat et al., 2007). A striking lack of verbal fluency is a frequent finding in PSP, and can be especially helpful in distinguishing PSP from idiopathic Parkinson’s disease (Rittman et al., 2013). Structural imaging can be helpful in the appropriate clinical context, with the characteristic findings of the ‘Mickey Mouse’ and ‘Morning Glory’ signs on axial section (Stamelou et al., 2011) and ‘Hummingbird’ or ‘Penguin’ sign on sagittal section (Kato et al., 2003, Massey et al., 2012) (Figure 1-7) being formalised in measurement of the midbrain to pons ratio (Massey et al., 2013).
Aetiology

PSP is caused by an accumulation of tau in midbrain structures (periaqueductal grey matter, red nucleus, colliculi, substantia nigra), dentate nucleus of the cerebellum and basal ganglia (Ellison et al., 2012). In the Richardson’s syndrome, Tau spreads to cortical brain regions only in advanced stages of the disease (Williams et al., 2007). In contrast to Alzheimer’s disease, the pathological tau deposits are composed of straight filaments of predominantly 4R tau (Taniguchi-Watanabe et al., 2016). The clinical variants of PSP share similar cellular and molecular features, but with differential distributions: i.e. presentation with focal cortical syndromes generally reflects greater cortical tau pathology, and with levodopa-responsive Parkinsonism reflects greater basal ganglia tau pathology (Dickson et al., 2010). Even within the Richardson syndrome, the distribution of disease burden predicts the severity of separable symptom domains (Ghosh et al., 2012).

1.4.4 Corticobasal Syndrome

Vignette
A 69 year-old professional organist attends clinic with a friend. For a year or more he has noticed difficulties with left sided coordination, which are especially noticeable when playing the organ. He finds that his left hand is less confident and his left foot is hesitant when trying to find the pedals. He has also had some difficulties with walking, and generally feels that his left leg is less responsive, but he has never fallen and has no true mobility problems. While he can still sight read 3 staves of music, he complains of difficulty integrating this into a single percept. In fact, he initially thought that his difficulties were with his eyes, but has now realised there is more going on. Despite this, he is still functioning at an extremely high level and even the choir master at his church has not noticed a problem with his playing. Outside of these domains he has very few problems. He has no REM sleep behaviour disorder. He has been hyposmic since a sinus washout 15 years ago but there has been no particular change. He describes normal forgetting such as going into rooms and forgetting the purpose, but also admits that he is a bit more reliant on his wife to remember appointments. On examination he has some left sided difficulties with both ideomotor and gesture copy praxis, and struggles with graphesthesia. Foot tapping is normal and there is no Parkinsonism. He scores 95/100 on ACE-III (attention 18/18, memory 23/26, fluency 13/14, language 25/26, visuospatial 16/16). Given the subtlety of his presentation, detailed neuropsychological assessment was requested, which revealed impairment across all cognitive domains, dominated by moderate memory and hand-sequencing deficits, with patchy performance across language, visuo-spatial and executive functioning tasks. His MRI demonstrated subtle biparietal atrophy, and a SPECT scan demonstrated neuronal hypometabolism in these areas (Figure 1-8). His dopamine transporter scan (DaT) demonstrated asymmetrically reduced tracer uptake (Figure 1-9).
Figure 1-8 MRI and SPECT images for the Corticobasal Syndrome vignette. Radiological convention. T1 weighted sagittal and coronal projections are shown, next to SPECT images from roughly the same projection. Moderate selective atrophy of parietal lobes can be seen, with a matching reduction in SPECT activity.

Figure 1-9 DaT scan summary image for the Corticobasal Syndrome vignette. Radiological convention. Reduced transporter activity can be seen, more marked on the right (in keeping with his predominantly left sided symptoms).
Clinical features and aetiology

It is important to distinguish corticobasal syndrome (CBS), which is a clinical diagnosis based on the criteria below, and corticobasal degeneration (CBD), which is a histopathological diagnosis based on a cortical distribution of tau lesions together with ballooned achromatic neurones and astrocytic plaques (Dickson et al., 2002). There are similarities between the histopathological appearances of CBD and PSP, and some regard them as being opposite ends of a spectrum of cortical vs brainstem involvement (Armstrong et al., 2013). The predictive value of CBS for a post-mortem diagnosis of CBD is poor (Shelley et al., 2009, Ling et al., 2010), and only a small proportion of pathologically confirmed CBD displayed CBS in life (Ling et al., 2010).

A ‘possible’ diagnosis of corticobasal syndrome can be made in the presence of at least one of the following: limb rigidity or akinesia; limb dystonia; limb myoclonus; and at least one of: orobuccal or limb apraxia; cortical sensory deficit; alien limb phenomena beyond simple levitation. The diagnosis is upgraded to ‘probable’ if the presentation is asymmetrical and at least two features from each category are present (Armstrong et al., 2013).

1.4.5 Multiple System Atrophy - MSA

Dementia is rarer in MSA than in the other neurodegenerative movement disorders, perhaps reflecting a younger age of onset; severe cognitive decline at presentation should prompt a reconsideration of the diagnosis. Addenbrooke’s Cognitive Examination score is usually in the normal range, especially if verbal fluency is excluded (Cope et al., 2014a). Characteristic cognitive deficits can, however, be revealed by detailed neuropsychological testing. Performance is impaired on tests of visuospatial and constructive function (Kawai et al., 2008), executive function (Robbins et al., 1992) and time perception (Cope et al., 2014a), especially in the Parkinsonian phenotype (MSA-P). Focussed intervention for these cognitive deficits can be helpful, for example the impairment of ability to internally generate a regular beat (Grahn and Rowe, 2012) means that patients with MSA have particularly significant benefits from the use of a metronome to pace their gait (Thaut et al., 1996) and prevent freezing (Okuma, 2006).
1.4.6 Huntington’s Disease - HD

**Vignette**

A 56 year old lady presents to clinic with progressively worsening slurred speech and difficulty with co-ordinating fine hand movements. Tasks like doing buttons, combing her hair and doing up her bra have become very difficult. Her balance has become poorer, but she has not fallen. Her husband mentions that she has begun to cram food into her mouth, earning the nickname ‘hamster’ from family members. You notice choreiform movements, but the patient says that these are not bothersome. Her mood is low, she is visibly anxious, and she becomes tearful while discussing her difficulties. Her maternal grandmother was said to have died from an aggressive form of Parkinson’s disease with dementia, while her mother died of cancer in her 40s. On examination you elicit a ‘milkmaid grip’, an unstable tandem gait, dysdiadochokinesis, symmetrical difficulties with Luria hand movements and finger tapping, impersistence of tongue protrusion and the initiation of saccades with a head thrust. ACE-III is 82/100 (attention 16/18, memory 19/26, fluency 7/14, language 24/26, visuospatial 14/16).

**Clinical Features**

While the chorea that characterises Huntington’s disease is immediately obvious to the observer, it is usually the least disabling of the symptoms. The neuropsychiatric features of HD are significant, and can be broadly separated into dementia, executive dysfunction, and affective disorder.

It is a fallacy that the motor disorder develops in advance of the dementia in HD. Even in cohorts with early disease specifically selected for seemingly normal cognition, significant cognitive deficits are revealed even by screening neuropsychology (Cope et al., 2014a). Detailed neuropsychological testing reveals difficulty on tests of executive function, planning, cognitive flexibility, abstract thought, action selection and time perception (Montoya et al., 2006, Tabrizi et al., 2009, Cope et al., 2014a). The affective disorder in HD is heterogeneous. Anxiety and irritability are the most common features at presentation often in conjunction with depression, but aggression, compulsive or addictive behaviours and sometimes inappropriate happiness are observed (Montoya et al., 2006, Klöppel et al., 2010). Psychosis does occur but, in the
absence of genetic confirmation of HD, its presence should prompt consideration of the HD mimics DRPLA (dentatorubral-pallidoluysian atrophy) (Adachi et al., 2001) and C9orf72 related neurodegeneration (Moss et al., 2014). An autosomal dominant family history of a movement disorder or dementia is usually present but, as here, misdiagnosis in earlier generations is common. At the single subject level structural imaging can appear normal in early HD, with the most commonly observed abnormality being flattening of the caudate head (Lang, 1985). The diagnostic genetic test is for an expansion of CAG triplet repeats in the gene encoding developmental protein Huntingtin, which lies on the short arm of chromosome 4. Fewer than 26 repeats is normal, and more than 40 predicts full penetrance of HD (Walker, 2007).

1.4.7 Normal Pressure Hydrocephalus - NPH

NPH is characterised by an insidious onset of dementia with psychomotor retardation and an apraxic gait, often associated with urinary incontinence, in the presence of ventricular enlargement and normal CSF pressure (Adams et al., 1965). It is a rare condition in younger individuals, but prevalence rises rapidly after the age of 80 (Jaraj et al., 2014). Ventricular enlargement has classically been defined as a ratio of greater than 0.3 between the maximum axial width of the frontal horns of the lateral ventricles and maximal internal diameter of skull at the same level (the so-called Evans’ index). This is a sensitive measure that lacks specificity, being present in 20% of elderly individuals (Jaraj et al., 2014) and an even greater proportion of those with neurodegenerative disease (Missori et al., 2016). Novel metrics such as the ratio between ventricular and intracranial volume (Toma et al., 2011) and a reduction in the angle of the corpus callosum on coronal section (Ishii et al., 2008) have increased specificity, and might have a role in case selection for neurosurgical management (Virhammar et al., 2014). NPH is, in many cases, treatable by one of a number of neurosurgical procedures that provide constant drainage of CSF, but case selection remains difficult (Anderson, 1986, Freimann et al., 2012). The gait disturbance commonly responds better than the cognitive impairment. At post mortem, NPH is strongly associated with demyelination of frontal regions in a pattern typical for small vessel disease (Akai et al., 1987). It remains unclear whether this white matter
pathology causes the altered CSF dynamics or is a consequence of altered periventricular metabolism (Jeppsson et al., 2013).

1.4.8 Niemann-Pick type C

Niemann-Pick disease types A and B are caused by autosomal recessive mutations in SMPD1 (Schuchman et al., 1992), which cause deficiency or absence of the lysosomal enzyme acid sphingomyelinase (Brady et al., 1966). This leads to a lipid storage disorder characterised by an accumulation of sphingomyelin (Klenk, 1935). Niemann-Pick type A is a severe, neurodegenerative disease of infancy that leads to death in the first few years of life, while type B is milder, causing hepatosplenomegaly, growth retardation and lung dysfunction, but generally exhibiting no neurological features.

Niemann-Pick type C is caused by autosomal recessive mutations in NPC1 (95% of cases) or NPC2, which respectively encode a large endosomal glycoprotein and a small, cholesterol-binding lysosomal protein (Ioannou, 2000). This leads to dysfunctional intracellular transport of cholesterol, leading to its sequestration in lysosomes and/or late endosomes. Age of onset is variable and, while approximately 50% of cases manifest by age 5, adult onset cases make up at least 20% of the total (Vanier and Millat, 2003). Mean age of onset in this subgroup is approximately 25 years, with the oldest reported case presenting at 54 (Vanier et al., 1991). Presentation is protean, but a progressive dementia is present in approximately 60% of cases. Other clinical features include cerebellar ataxia (76%), vertical supranuclear ophthalmoplegia (75%), dysarthria, (63%), movement disorders (58%), splenomegaly (54%), psychiatric disorders (45%) and dysphagia (37%) (Sevin et al., 2007). Epilepsy and cataplexy are more rarely reported, but the presence of gelastic cataplexy in the absence of a sleep disorder is highly suggestive of Niemann-Pick type C.

Psychiatric disturbance is the presenting feature in more than a third of cases, and can be manifest for a number of years before visceral or deep brain signs appear. Psychosis is the most common manifestation, but bipolar disorder, isolated depression and obsessive compulsive disorder have been reported (Sullivan et al., 2005). While onset can be progressive, most patients present acutely and have relapsing and remitting disease, leading to an initial diagnosis of schizophrenia before additional clinical
features emerge. It is rare but recognised for psychiatric disturbance to complicate later disease in patients initially presenting with motor features.

About a quarter of patients present with isolated cognitive dysfunction. Again, the phenotype is variable, from mild dysexecutive features to aggressive global dementia. Interestingly, neurofibrillary tangles and beta amyloid plaques have been observed in relatively young patients with Niemann-Pick type C (Saito et al., 2002), but their presence is the exception not the rule.

Diagnostic testing for Niemann-Pick type C is complex. Classically, the diagnostic cornerstone has been skin biopsy with fibroblast culture and the demonstration of impaired intracellular cholesterol transport with filipin staining (Patterson et al., 2012). This test is strongly positive in about 80% of cases, and can also give a clear negative result, but about 20% of patients show only intermediate staining. In recent years this has been complemented by next generation sequencing of the NPC1 and NPC2 genes, but novel mutations of uncertain significance are commonly found and interpretation can be difficult. Simpler tests, based on mass spectroscopy of plasma biomarkers, have recently become available; validation studies are ongoing, but initial evidence suggests that these tests might prove sensitive but not specific (Vanier et al., 2016).

Disease modification can be achieved with Miglustat, a small molecule inhibitor of glucosylceramide synthase (Patterson et al., 2007) that has been demonstrated to reduce the rate of disease progression. Current clinical consensus is that all patients with neurological, psychiatric or cognitive manifestations should be offered this medication. The primary side effects are gastrointestinal, and can be partially ameliorated with dietary and anti-propulsive strategies (Belmatoug et al., 2011). There is also a significant role for symptomatic therapy to treat seizures, cataplexy, dystonia, psychiatric disturbance and sleep disorders.

1.4.9 Spinocerebellar Ataxia - SCA

SCA is a collection of more than 60 genetic neurodegenerative diseases that are united by their primary manifestation with ataxia and (usually) cerebellar degeneration, but with different associated features. Most are autosomal dominant with adult onset, but X-linked and childhood onset forms are recognised (Stephen et al., 2016). The SCAs
are assigned a number according to the date of discovery of the associated gene, and the classification is consequently confusing and difficult to remember (see Worth, 2004 for a clinically focussed review). While even those disorders such as SCA-6 thought to result in ‘pure’ cerebellar dysfunction display a subtle ‘cerebellar cognitive phenotype’ of impairments in cognitive flexibility, response inhibition, verbal reasoning (Cooper et al., 2010) and time interval perception (Grube et al., 2010), many of the SCAs display a broader cognitive phenotype. Cognitive impairment, hallucinations or depression are especially common as an early or presenting feature in SCA-17, which is caused by a triplet repeat expansion in TBP (Rolfs et al., 2003).

1.5 Cognitive impairments that may or may not progress

In this section I introduce the concepts of mild cognitive impairment and subjective cognitive impairment. These diagnostic labels are agnostic as to the underlying aetiology of the impairment, and as such they may or may not progress. I then conclude the chapter with brief descriptions of two conditions that can present to the cognitive clinic but that do not reflect neurodegenerative disease; transient epileptic amnesia and transient global amnesia.

1.5.1 Mild Cognitive Impairment

Vignette

A 62 year old mathematics teacher presents to clinic alone. She complains that her memory is generally poorer than it was, and she is finding it especially difficult to learn the names of children in her class. She is not struggling with the mathematical aspect of their work, but has found it difficult to adapt to curriculum changes and new information technology. Her mood is stable, and she denies any significant stressors. ACE-III is 89 (attention 16/18, memory 20/26, fluency 13/14, language 26/26, visuospatial 16/16).

Clinical Features

The primary differentiation between dementia and mild cognitive impairment (MCI) is one of degree. Generally speaking, a diagnosis of MCI is made when deficits are
insufficiently severe to interfere with function at work or in daily activities, but this is not a firm criterion. The distinction is a sensitive clinical judgment that requires a skilled clinician to make an holistic assessment of the patient from primary and secondary sources. In the vignette above, the patient’s work function was affected, but a similar level of difficulty might not have had such a noticeable effect on a different profession or in one retired. Clearly it is not intellectually satisfying for the same constellation of symptoms and signs to lead to different diagnoses dependent solely on work status. Another factor that must weigh on this decision is the level of certainty that one is dealing with a progressive condition. A diagnosis of dementia carries with it the inevitable long-term implications of a neurodegenerative condition; the same is not necessarily true of MCI, from which only a proportion of individuals will progress. Some clinicians might find it appropriate in mild cases to make an initial diagnosis of MCI pending imaging and longitudinal neuropsychological assessment, while others prefer to refrain from making any diagnosis until progression has been assessed – neither approach is incorrect, and this decision should be made on an individual basis in a patient-centred manner.

Management

Once structural and significant vascular pathology has been excluded with structural imaging, the most important management step in MCI is to closely monitor for the emergence of a dementia requiring medical or social intervention. It is possible to subclassify mild cognitive impairment by the cognitive domain or domains affected. It has been proposed by some that those with a predominantly amnestic phenotype are more likely to progress to Alzheimer’s disease, and that this group should be further investigated with biomarkers of amyloid deposition and neuronal injury such as CSF tau/phosphorylated tau and Aβ42, PET amyloid imaging (PiB or Florbetapir), or perfusion imaging with FDG-PET or SPECT (Albert et al., 2011). Others have gone further, suggesting that amnestic MCI with supportive biomarkers should be termed prodromal Alzheimer’s disease (Dubois and Albert, 2004). While of undoubted utility in research, where the current focus of therapeutic trials is in early disease, this approach is clinically available in a limited number of centres and is not generally advocated by the UK National Institute for Clinical Excellence (NICE and SCIE, 2006).
Clearly the emergence of disease modifying therapeutics will change the balance of this decision.

1.5.2 Subjective Cognitive Impairment

Vignette

A 52 year old gentleman attends clinic alone. He left school at 16 with three CSEs, and became a planning officer at a telecoms company, where he undertook an apprenticeship and several postgraduate qualifications. He describes several years of poor anterograde memory. He has developed elaborate coping mechanisms such as littering his house with post-it notes and ‘quick guides’ for new computer programs. He reports poor day-to-day memory, but is unable to give concrete examples of this. He presents due to a perceived worsening in symptoms. There have been no concerns from his partner or employer about performance. He reports anxiety at work, but ascribes this to his perceived performance difficulties; there is no overt mood disorder. Neurological examination is normal. ACE-III is 96 (attention 18/18, memory 23/26, fluency 13/14, language 26/26, visuospatial 16/16). Detailed neuropsychological assessment is undertaken, which reveals normal performance on tests of verbal comprehension, perceptual reasoning, executive function, processing speed and language, but catastrophic failure on tests of memory. The pattern of failure is a lack of immediate recall of both a story and the Rey complex figure, with consequent impairment of delayed recall. The neuropsychologist reports that he displayed significant anxiety during testing. Screening blood tests were normal, as was structural brain scanning.

Clinical Features

A subjective cognitive impairment is defined as the presence of cognitive symptoms in the absence of objective evidence of cognitive impairment or psychopathy (Reisberg et al., 2010, Stewart, 2012). It is characterised by a seeming inconsistency between disability and measureable deficit. While it has traditionally been viewed as a syndrome of hyper-vigilance, there is now converging evidence that at the population level it is a predictor of risk of future decline. Those with subjective cognitive impairment do show small differences from controls on neuropsychology and imaging at the group level.
(Hohman et al., 2011), and subjective cognitive impairment sufferers over the age of 65 have a hazard ratio of 2.3 for acquiring dementia within four years (Waldorff et al., 2012). Nonetheless, the appropriate clinical management is usually reassurance and practical suggestions such as list keeping and sleep hygiene. Where it is evident, anxiety should be addressed.

1.5.3 Transient Global Amnesia - TGA

TGA is a syndrome characterised by an isolated temporary amnesia for a period of a few hours (Fisher, 1958). During an attack autobiographical memory is intact, but patients are frequently disoriented in time, place and person, and ask repetitive questions on these topics (Bolwig, 1968). After an attack, there is a total loss of episodic memory for the period of confusion and, usually, a period of time beforehand. The aetiology is unknown. It has been observed that vascular risk factors are over-represented in patients with TGA compared to controls (Shuping et al., 1980a), but under-represented compared to patients with TIA (Melo et al., 1992). Some individuals display diffusion-weighted imaging changes in medial temporal regions during and after an attack, but the evidence for an ischaemic cause is somewhat conflicted (see (Huber et al., 2002) for a review). The syndrome is usually monophasic, with approximately 6% of patients experiencing a second attack (Melo et al., 1992). Recurrent attacks should prompt a search for a structural (Shuping et al., 1980b) or embolic (Shuping et al., 1980a) cause.

1.5.4 Transient Epileptic Amnesia - TEA

TEA is a syndrome of recurrent, brief attacks of amnesia due to focal temporal lobe seizures. It characteristically occurs on waking, and is most common in middle-age (Zeman and Butler, 2010). Like TGA, repetitive questioning is common, but recurrent attacks are characteristic and amnesia for the attack is sometimes less dense (Zeman et al., 1998). Important distinguishing symptoms are a loss of remote autobiographical knowledge (Butler et al., 2007) and accelerated long-term forgetting (Muhlert et al., 2010). The latter is a fascinating symptom, in which memories are initially well encoded but fade more rapidly than normal over subsequent days. The wonderfully
eccentric defining study of this phenomenon involved patients and their spouses or friends visiting the stately homes of Britain while wearing video cameras. When quizzed on the same day about what they had seen, patients performed as well as controls, but recall across a number of modalities (images, thoughts, events and sensory information) was much poorer after a delay of days to weeks. While the seizures, and consequently the discrete amnestic attacks, often respond well to small doses of anticonvulsant medication, accelerated long-term forgetting frequently persists (Butler, 2006).

1.6 Challenges and opportunities

The ultimate goal of dementia research is to prevent or reverse deteriorations in cognitive function, early in the evolution of disease. The diverse array of clinical syndromes that comprise the dementias, and their differing underlying causative pathologies, represent a significant challenge to this goal. Therefore, the original research I outline in this thesis assesses specific neurodegenerative diseases as demonstrator conditions, informing us not only about the conditions themselves but allowing a more general insight into what might be the ultimate similarities and differences in approach between successful treatments for the conditions reviewed in this introductory chapter.

Neurodegenerative diseases are united by early effects in synaptic loss, reduced plasticity and aberrant connectivity that precede cell death and atrophy. However, patterns of connectivity in the healthy brain have strong influences on the vulnerability of brain regions to the accrual of neuropathology. It is therefore often difficult to dissociate the causes of neurodegeneration from its consequences. In Chapter 2, I address this question in two diseases both characterised by abnormal intracellular deposits of hyper-phosphorylated filamentous tau inclusions but with very different distributions. I assess three competing hypotheses of neuronal vulnerability to explain why tau in Alzheimer’s disease impacts large-scale cortical connectivity networks but, by contrast, tau in PSP is confined to midbrain and deep brain nuclei in early to mid-stage disease. Understanding the variation in tau expression is an essential step to
developing appropriate treatments to prevent its accrual, and hopefully reduce its harmful effects, in differing clinical syndromes.

As well as understanding the pathogenesis of neurodegeneration, the development and evaluation of therapeutic interventions at economically viable timescales will require validated models of the neurocognitive code and the tools to assess it in patients. I chapter 2 I assess the effects of regional neurodegeneration on the observable functional connectivity of the brain in terms of correlated activity. In chapter 3, I go beyond this by directly observing the consequences of regional neurodegeneration on brain activity in unaffected regions. To do this I use the frontal aphasic form of frontotemporal dementia (nfvPPA) as demonstrator condition, because of its stereotyped pattern of neurodegeneration, the psychophysical precision of language as a probe, and the fragility of human language networks to synaptic disruption. By characterising the neurophysiology of this disease using magnetoencephalography (MEG), and explaining patients’ behaviour with Bayesian computational models, I provide causal evidence about the mechanisms of normal human higher cognitive functioning and neural information coding. This leads on to chapter 4, in which I undertake a detailed psychophysical examination of the brain basis of agrammatism in nfvPPA, as a model of higher order cognitive disruption in neurodegeneration, and discuss the potential utility of therapeutic approaches to symptom rehabilitation.

Overall, this use of demonstrator conditions within the broader clinical context of dementia allows for the use of systems neuroscience as a translational bridge between the molecular biology of dementia and targeted clinical trials. Further, it improves our understanding of normative brain networks through the study of focal forms of neurodegeneration, as the basis for new frameworks for cognitive deficits in neurodegenerative disease. Together, this approach holds the promise of translating recent fantastic progress in understanding the cellular mechanisms of neurodegeneration, and success in rescuing animal models, to the development of effective disease modifying therapies for dementia in humans.
1.7 Chapter Summary:

The dementias are persistent or progressive disorders affecting more than one cognitive domain that interfere with an individual’s ability to function at work or home, and represent a decline from a previous level of function. The assessment of dementias is a multi-disciplinary pursuit, with clinical and diagnostic assessments undertaken by psychiatrists, psychologists, neurologists, geriatricians and, increasingly, specialist nurses, occupational therapists and other allied health professionals.

The dementias can be classified either by their cognitive phenotype, or by their aetiological origin, and there is not necessarily a linear mapping between the two. This chapter outlines the neurodegenerative causes of cognitive impairment in terms of their clinical features, epidemiology and aetiology. For each I provide a brief discussion of management and prognosis, focussing on the differences between the syndromes rather than providing a general approach to the patient with a memory complaint or neurodegenerative cognitive impairment.

The intention of this introductory chapter is to set the scene for the experimental chapters that follow, in which I undertake in-depth studies of particular aspects of the physiology of dementia. For this, I will use model diseases that allow controlled experiment and robust comparison. These diseases serve as examples of how cellular and molecular pathologies give rise to aberrant physiological function across widespread brain networks, which in turn leads to deficits in core cognitive systems, resulting in the complex symptomatology of individual dementia syndromes.
Chapter 2: Tau burden and the functional connectome in Alzheimer's disease and progressive supranuclear palsy

2.1 Preface

In this experimental chapter I assess the relationship between functional connectivity and \textit{in vivo} tau burden by combining graph theoretical analyses of resting state functional MR imaging and PET imaging using the ligand AV-1451. I do this in two model diseases, Alzheimer’s disease and Progressive Supranuclear Palsy, which represent a contrast between predominantly ‘cortical’ and ‘subcortical’ burdens of tau neuropathology.

A paper based on this chapter has been published in \textit{Brain} with myself as first author (Cope \textit{et al.}, 2018). The original data on which the analyses in this chapter were based were collected by Ms Patricia Vazquez-Rodriguez, Dr Richard Bevan-Jones and Mr Robert Arnold; patients were scanned and assessed as part of the NIMROD study (Bevan-Jones \textit{et al.}, 2017b), which was conceptualised by Prof John O’Brien and Prof James Rowe. The neuroimaging data were pre-processed by Mr Robin Borchert and Mr P Simon Jones. I am particularly grateful to Dr Tim Rittman for introducing me to graph metrics, and for using the software he developed during his own PhD to derive the graph theoretical measures from the functional connectivity data (Maybrain: github.com/rittman/maybrain). I am grateful to Dr Luca Passamonti and Dr Deniz Vatansever for helpful comments on interpretation, and to Dr Kieren Allinson for his expertise on \textit{post mortem} neuropathological distribution. Finally, I am grateful to three anonymous reviewers at \textit{Brain}, who suggested the useful addition of several additional graph measures. I designed the analysis strategy, performed the analyses, constructed all of the figures, and wrote the text with comments from all of those listed above.
2.2 Introduction

Alzheimer’s disease and Progressive Supranuclear Palsy (PSP) are both characterised by intracellular neurofibrillary lesions containing hyper-phosphorylated filamentous tau inclusions (Goedert and Spillantini, 2006). However, the diseases differ in: 1) the distribution of these tau inclusions; 2) the balance of expression of tau isoforms; and 3) the ultrastructure of tau filaments. Here, I test the impact of these differences in tau pathology on the reorganisation of large-scale functional brain connectivity architecture.

Alzheimer’s disease is characterised by widespread extracellular deposition of amyloid-β and paired helical filaments of tau with three (3R) and four (4R) repeats in the microtubule-binding domain (Sisodia et al., 1990, Liu et al., 2001). In Alzheimer’s disease, these pathological proteins arise early in the transentorhinal cortex, from where they spread to limbic regions, followed by inferior frontal and parietal cortex (Braak and Braak, 1991, Braak and Braak, 1995). Although tau and amyloid-β have complex and synergistic effects (Nisbet et al., 2015), there is converging evidence that tau mediates direct toxic effects on neurons and synaptic plasticity (Ballatore et al., 2007, Roberson et al., 2007, Ittner and Götz, 2011, Myeku et al., 2016) and correlates with hypo-metabolism and symptomatology in Alzheimer’s disease (Lehmann et al., 2013, Ossenkoppele et al., 2015, Ossenkoppele et al., 2016). By contrast, pathological tau deposits in PSP are composed of straight filaments of predominantly 4R tau (Taniguchi-Watanabe et al., 2016). This is most prominent in midbrain and deep brain nuclei early in the course of the Richardson’s syndrome variant of PSP, spreading to cortical regions in advanced stages of the disease (Williams et al., 2007).

To estimate the burden of tau pathology in vivo, PET ligands have been developed, including $[^{18}F]$AV-1451 (Chien et al., 2013, Xia et al., 2013, Ossenkoppele et al., 2015, Ossenkoppele et al., 2016). While there has been significant debate about the specificity of this ligand for tau, especially in the non-Alzheimer’s dementias (Bevan-Jones et al., 2017a, Xia and Dickerson, 2017), the distribution of $[^{18}F]$AV-1451
binding recapitulates Braak staging in Alzheimer’s disease (Schwarz et al., 2016) and correlates with post-mortem neuropathology in primary tauopathies (Smith et al., 2016). [\textsuperscript{18}F]AV-1451 binding has been shown to correlate with cognitive performance in Alzheimer’s disease (Johnson et al., 2016) and healthy older adults (Schöll et al., 2016) more robustly than quantitative beta amyloid imaging (Brier et al., 2016). The [\textsuperscript{18}F]AV-1451 ligand also effectively distinguishes between AD and PSP cases on the basis of both the intensity and regional distribution of its binding potential (BP\textsubscript{ND}) (Passamonti et al., 2017).

In many neurodegenerative disorders, neuropathology and atrophy are most marked in those brain regions that are densely connected, both at the structural (Crossley et al., 2014) and functional level (Dai et al., 2014). In graph-theoretical terms, these densely connected regions are usually referred as ‘hubs’ (Buckner et al., 2009). There are a number of hypotheses as to why hubs are vulnerable to neurodegeneration. First, pathological proteins may propagate trans-neuronally, in a prion-like manner (Prusiner, 1984, Baker et al., 1994, Goedert, 2015) such that highly connected regions are more likely to receive pathology from ‘seed’ regions affected in early stages of the disease (Zhou et al., 2012), leading to neurodegeneration that mirrors structural and functional brain connectivity (Raj et al., 2012, Abdelnour et al., 2014, Raj et al., 2015). Alternatively, hubs might be selectively vulnerable to a given level of pathology, due to a lack of local trophic factors (Appel, 1981), higher metabolic demands (Saxena and Caroni, 2011, de Haan et al., 2012), or differential gene expression (Rittman et al., 2016).

These alternate hypotheses lead to different predictions about the relationship between tau burden and connectivity. The trans-neuronal spread hypothesis predicts that regions that are more strongly interconnected would accrue more tau pathology. This would manifest as higher tau burden in nodes with larger weighted degree, which is a measure of the number and strength of functional connections involving each node. In contrast, if hubs are vulnerable to tau accumulation because of increased metabolic demand this might manifest as a positive relationship between tau burden and
participation coefficient, which is a measure of the proportion of a node’s connections that are with other neural communities, and is the graph metric that is most closely correlated with metabolic activity (Chennu et al., 2017). Finally, if trophic support is an important factor in tau accumulation, this might manifest as a negative relationship between tau burden and clustering coefficient; nodes with less tightly clustered connectivity patterns might have more vulnerable trophic supply.

Here I go beyond previous associative studies to examine, in the same subjects, the relationship between in vivo tau burden, as measured by the PET ligand $[^{18}F]$AV-1451 BP$_{ND}$, and functional connectivity, as summarised by graph theoretic measures based on resting state (task-free) functional magnetic resonance imaging (fMRI) (Sporns et al., 2004, Bullmore and Sporns, 2009, Rubinov and Sporns, 2010). I test the following linked hypotheses:

1) Brain regions that are normally more densely interconnected accrue more tau pathology.

2) In Alzheimer’s disease, where tau accumulation is predominantly cortical, the functional consequence is that affected nodes become more weakly connected and local efficiency of information transfer is reduced.

3) In PSP-Richardson’s syndrome, where neurodegeneration associated with tau accumulation is most severe in midbrain and basal ganglia, the functional consequence of disrupted cortico-subcortical and cortico-brainstem interactions (Gardner et al., 2013) is that indirect cortico-cortical connections become stronger.

2.3 Methods

2.3.1 Participants

All patients had mental capacity to take part in the study and provided informed consent. Study procedures were approved by the National Research Ethics Service. I recruited 17 patients with Alzheimer’s disease, as evidenced by either a clinical diagnosis of probable Alzheimer’s dementia according to consensus criteria ($n = 10$)
(McKhann et al., 2011), or a clinical diagnosis of mild cognitive impairment (MCI) and a positive amyloid PET scan \((n = 7)\) (Klunk et al., 2004, Okello et al., 2009). Patients with MCI and evidence of beta amyloid were included to ensure the largest possible spread in tau burden within the Alzheimer’s disease group. I also recruited 17 patients with PSP-Richardson’s syndrome by 1996 criteria (Litvan et al., 1996). Retrospective case review confirmed that all PSP subjects also met the revised 2017 criteria for PSP-RS (Höglinger et al., 2017). 12 age matched controls were also examined. Participant demographics are shown in Table 2-1 and detailed neuropsychological test results for each subject are shown in Table 2-2.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Age</th>
<th>Years of Education</th>
<th>MMSE</th>
<th>ACE-R</th>
<th>PSP-RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>17</td>
<td>71 (9)</td>
<td>14 (3)</td>
<td>25 (4)</td>
<td>72 (14)</td>
<td>-</td>
</tr>
<tr>
<td>PSP</td>
<td>17</td>
<td>69 (6)</td>
<td>12 (2)</td>
<td>27 (4)</td>
<td>83 (14)</td>
<td>41 (14)</td>
</tr>
<tr>
<td>Controls</td>
<td>12</td>
<td>67 (8)</td>
<td>16 (2)</td>
<td>29 (1)</td>
<td>96 (3)</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2-1 Participant demographics summary

<table>
<thead>
<tr>
<th>Group</th>
<th>ACE-R</th>
<th>MMSE</th>
<th>BADLS</th>
<th>CAF</th>
<th>CBI-R</th>
<th>GDS</th>
<th>HADS</th>
<th>INECO</th>
<th>NPI</th>
<th>PAL</th>
<th>PPT</th>
<th>RAVLT</th>
<th>SRT</th>
<th>UPDRS</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>73</td>
<td>24</td>
<td>6</td>
<td>2</td>
<td>23</td>
<td>0</td>
<td>3</td>
<td>16</td>
<td>12</td>
<td>49</td>
<td>43</td>
<td>23</td>
<td>374</td>
<td>4</td>
</tr>
<tr>
<td>AD</td>
<td>57</td>
<td>21</td>
<td>3</td>
<td>2</td>
<td>38</td>
<td>0</td>
<td>6</td>
<td>19</td>
<td>5</td>
<td>7</td>
<td>49</td>
<td>11</td>
<td>311</td>
<td>0</td>
</tr>
<tr>
<td>AD</td>
<td>50</td>
<td>17</td>
<td>11</td>
<td>2</td>
<td>94</td>
<td>2</td>
<td>3</td>
<td>12</td>
<td>51</td>
<td>71</td>
<td>50</td>
<td>11</td>
<td>499</td>
<td>1</td>
</tr>
<tr>
<td>AD</td>
<td>41</td>
<td>15</td>
<td>16</td>
<td>1</td>
<td>59</td>
<td>5</td>
<td>6</td>
<td>22</td>
<td>N/A</td>
<td>N/A</td>
<td>14</td>
<td>748</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AD</td>
<td>82</td>
<td>26</td>
<td>5</td>
<td>0</td>
<td>38</td>
<td>1</td>
<td>4</td>
<td>18</td>
<td>11</td>
<td>11</td>
<td>50</td>
<td>33</td>
<td>236</td>
<td>0</td>
</tr>
<tr>
<td>AD</td>
<td>75</td>
<td>25</td>
<td>10</td>
<td>7</td>
<td>66</td>
<td>9</td>
<td>15</td>
<td>9</td>
<td>26</td>
<td>14</td>
<td>48</td>
<td>8</td>
<td>335</td>
<td>0</td>
</tr>
<tr>
<td>AD</td>
<td>69</td>
<td>22</td>
<td>1</td>
<td>N/A</td>
<td>20</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>10</td>
<td>64</td>
<td>44</td>
<td>19</td>
<td>241</td>
<td>0</td>
</tr>
<tr>
<td>AD</td>
<td>70</td>
<td>27</td>
<td>17</td>
<td>5</td>
<td>93</td>
<td>1</td>
<td>1</td>
<td>19</td>
<td>32</td>
<td>40</td>
<td>48</td>
<td>15</td>
<td>305</td>
<td>13</td>
</tr>
<tr>
<td>AD</td>
<td>53</td>
<td>19</td>
<td>18</td>
<td>3</td>
<td>52</td>
<td>6</td>
<td>11</td>
<td>11</td>
<td>21</td>
<td>75</td>
<td>42</td>
<td>29</td>
<td>472</td>
<td>2</td>
</tr>
<tr>
<td>AD</td>
<td>60</td>
<td>22</td>
<td>10</td>
<td>0</td>
<td>39</td>
<td>0</td>
<td>4</td>
<td>12</td>
<td>6</td>
<td>49</td>
<td>49</td>
<td>19</td>
<td>417</td>
<td>27</td>
</tr>
<tr>
<td>AD</td>
<td>80</td>
<td>27</td>
<td>8</td>
<td>0</td>
<td>28</td>
<td>3</td>
<td>10</td>
<td>21</td>
<td>14</td>
<td>43</td>
<td>51</td>
<td>26</td>
<td>343</td>
<td>0</td>
</tr>
<tr>
<td>AD</td>
<td>85</td>
<td>28</td>
<td>2</td>
<td>2</td>
<td>18</td>
<td>1</td>
<td>2</td>
<td>15</td>
<td>0</td>
<td>48</td>
<td>50</td>
<td>30</td>
<td>216</td>
<td>0</td>
</tr>
<tr>
<td>AD</td>
<td>84</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>3</td>
<td>17</td>
<td>18.5</td>
<td>0</td>
<td>46</td>
<td>51</td>
<td>26</td>
<td>235</td>
<td>0</td>
</tr>
<tr>
<td>AD</td>
<td>83</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>3</td>
<td>51</td>
<td>35</td>
<td>346</td>
<td>1</td>
</tr>
<tr>
<td>AD</td>
<td>72</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>2</td>
<td>6</td>
<td>15</td>
<td>4</td>
<td>13</td>
<td>42</td>
<td>20</td>
<td>365</td>
<td>1</td>
</tr>
<tr>
<td>AD</td>
<td>77</td>
<td>23</td>
<td>2</td>
<td>4</td>
<td>18</td>
<td>0</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>45</td>
<td>47</td>
<td>18</td>
<td>363</td>
<td>2</td>
</tr>
</tbody>
</table>

49
### Table 2.2 Detailed neuropsychological profiles of individual study participants.

| Control | 97  | 30  | N/A | N/A | N/A | 0  | 5  | 26.5 | N/A | 0  | 52  | 236 | 0  |
| Control | 98  | 30  | N/A | N/A | N/A | 1  | 6  | 24   | N/A | 2  | 52  | 45  | 322 | 0  |
| Control | 97  | 30  | N/A | N/A | N/A | 1  | 9  | 27   | N/A | 11 | 50  | 259 | 0  |         |
| Control | 98  | 30  | N/A | N/A | N/A | 3  | 4  | 21   | N/A | 4  | 52  | 33  | 233 | 0  |
| Control | 93  | 30  | N/A | N/A | N/A | 10 | 28 | 23.5 | N/A | 2  | 52  | 43  | 254 | 0  |
| Control | 98  | 28  | N/A | N/A | N/A | 5  | 21 | 27   | N/A | 6  | 50  | 256 | N/A |
| Control | 96  | 29  | N/A | N/A | N/A | 1  | 3  | 29   | N/A | 0  | 52  | 207 | 0  |         |
| Control | 99  | 30  | N/A | N/A | N/A | 2  | 8  | 26   | N/A | 6  | 51  | 36  | 232 | 0  |
| Control | 97  | 30  | N/A | N/A | N/A | 0  | 4  | 22.5 | N/A | 0  | 51  | 45  | 237 | N/A |
| Control | 95  | 29  | N/A | N/A | N/A | 0  | 2  | 23   | N/A | 4  | 52  | 45  | 238 | N/A |
| Control | 90  | 28  | N/A | N/A | N/A | 3  | 11 | 21.5 | N/A | 3  | 52  | 39  | 252 | N/A |

N/A = data not available. ACE-R: Addenbrooke's Cognitive Examination, Revised; MMSE: Mini Mental State Examination; BADLS: Bristol Activities of Daily Living; CAF: Clinical Assessment of Fluctuation; CBI-R: Cambridge Behavioural Inventory Revised; GDS: Geriatric Depression Scale; HADS: Hospital Anxiety and Depression Scale Combined Score; INECO: Institute of Cognitive Neurology Frontal Screening; NPI: Neuropsychiatric Inventory; PAL: Paired Associates Learning; PPT: Pyramids and Palm Trees; RAVLT: Rey Auditory Verbal Learning; SRT: Simple Reaction Time (mean); UPDRS: Unified Parkinson's Disease Rating Scale.
2.3.2 Study Procedures

This study formed part of the NIMROD (Neuroimaging of Inflammation in Memory and Related Other Disorders) project, for which the trial protocol containing general methods has been previously published (Bevan-Jones et al., 2017b).

All participants underwent the Addenbrooke’s Cognitive Examination-Revised (ACE-R), eleven minutes of resting state functional MRI at 3 Tesla, and $^{[18F]}$AV-1451 PET imaging. Participants with PSP were also examined according to the PSP-Rating Scale (Golbe and Ohman-Strickland, 2007). Those with a clinical diagnosis of MCI had $^{[11C]}$PiB PET imaging on a separate visit – only those with increased uptake indicative of underlying Alzheimer’s pathology are reported here.

2.3.3 MRI data acquisition and pre-processing

MR imaging was performed at the Wolfson Brain Imaging Centre, University of Cambridge, UK using a 3T Siemens Magnetom Tim Trio scanner with a Siemens 32-channel phased-array head coil (Siemens Healthcare, Erlangen, Germany).

A T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) image was acquired with repetition time (TR) = 2300ms, echo time (TE) = 2.98ms, matrix = 256x240, in-plane resolution of 1x1mm, 176 slices of 1mm thickness, inversion time = 900ms and flip angle = 9°.

Eyes-closed resting state (task-free) multi-echo functional imaging was carried out for eleven minutes. A total of 269 EPI image volumes were acquired with TR = 2430ms, TEs = 13.00, 30.55, and 48.10ms, matrix = 64x64, in-plane resolution of 3.75x3.75mm, 34 slices of 3.8mm thickness with an interslice gap of 0.38mm, GRAPPA parallel imaging with an acceleration factor of 2 and bandwidth = 2368 Hz/pixel. The first six volumes were discarded to eliminate saturation effects and achieve steady state magnetization. Preprocessing employed the ME-ICA pipeline.
(https://wiki.cam.ac.uk/bmuwiki/MEICA) (Kundu et al., 2012, Kundu et al., 2013), which uses independent component analysis to classify BOLD and non-BOLD signals based on the identification of linearly dependent and independent echo time (TE) dependent components. This provides an optimal approach to correct for movement-related and non-neuronal signals, and is therefore particularly well suited to our study, in which systematic differences in movement or head position might reasonably have been expected between patient groups. In fact, perhaps surprisingly, such differences were not observed (Figure 2-1) – between group ANOVAs demonstrated no significant differences between groups in terms of frame displacement before (p = 0.88) or after (p = 0.42) ME-ICA pre-processing, nor in terms of DVARS (Power et al., 2012) before (p = 0.89) or after (p = 0.67) ME-ICA pre-processing. Note that both movement parameters were approximately an order of magnitude lower after ME-ICA pre-processing.

Figure 2-1 Movement parameters. Movement quantified by frame displacement and DVARS (D referring to temporal derivative of timecourses, VARS referring to RMS variance over voxels), before and
after MEICA de-noising. No significant differences were demonstrated between groups.

The MPRAGE images were processed into the standard space using DARTEL (Ashburner, 2007), producing a study specific template in stereotactic space. Each fMRI series mean image was co-registered to the corresponding MPRAGE image. The whole fMRI series was warped to the template space using the DARTEL flow fields.

In order to perform a whole brain graph-theoretical analysis that included the brainstem, cerebellum and subcortical structures, the Harvard-Oxford Cortical atlas and Harvard-Oxford subcortical atlas, each thresholded at 25%, were combined. Additionally, a Freesurfer 6 brainstem parcellation of the MNI152 (2009 asymmetric) brain together with remaining Ventral DC completed the whole brain labelling. This atlas was sub-parcellated into 598 regions of approximately equal volume (mean 1.995, SD 0.323ml) such that each sub-parcel could be uniquely identified with an atlas region. The MNI-space parcellation was matched to the group standard space using inverse deformations following application of the ‘Population to ICBM Registration’ SPM function to the group template, with nearest neighbour interpolation.

The BOLD time series for each node was extracted using the CONN functional connectivity toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012). Between-node association matrices were generated, and then z-transformed for further analysis.

2.3.4 Functional Connectivity Assessment

Graph theoretical analysis was used to investigate the global and local characteristics of brain networks. These metrics were calculated in python using the Maybrain software (github.com/rittman/maybrain) and networkx (version 1.11). Reported metrics were calculated in each subject at absolute network density thresholds of 1-10% in 1% increments using a minimum spanning tree to ensure complete connectivity of the graph (Alexander-Bloch et al., 2010). Primary statistical analysis was performed at an intermediate density of 6%, with confirmatory analyses separately performed at all other densities. Weighted degree and participation coefficient were analysed in their
raw forms, and all other metrics were dissociated from variation in degree by binarisation after thresholding and normalisation against 1000 random graphs with the same number of connections at each node.

The graph metrics assessed were:

i) Weighted degree: the number and strength of functional connections involving each node.

ii) Weighted participation coefficient: the proportion of a node’s functional connectivity that involves other nodes that are not part of its own community structure, as defined by the Louvain community detection algorithm. Nodes with high degree and low participation coefficient are ‘provincial hubs’ (i.e. display strong connectivity only within their own community), while those with high participation coefficient are connector nodes (Joyce et al., 2010).

iii) Betweenness centrality: the number of shortest paths between any other two nodes that pass through the node of interest. Nodes that are important for the transfer of information between other nodes have high betweenness centrality.

iv) Closeness centrality: the inverse of the path length between a node and all other nodes in the graph. This is the node-wise equivalent of global efficiency, which is the inverse sum of all the shortest path lengths in the graph.

v) Local efficiency: the number of strong connections a node has with its neighbouring nodes. This reflects the robustness of local networks to disruption.

vi) Eigenvector centrality: this measure quantifies the functional influence of a node on every other node in the graph, by weighting the importance of each nodal connection based on the influence of the nodes with which they connect.

vii) Clustering coefficient: the fraction of triangular connections formed by a node with other nodes. In other words, a node is strongly clustered if a large proportion of its neighbours are neighbours of each other.

Nodal connectivity strength was assessed for comparison to weighted degree, to ensure that our results did not result from bias introduced by proportionate thresholding. This metric is related to weighted degree, but includes information from all strengths of
connection between every pair of nodes. As such, it is more subject to fMRI signal to noise ratio limitations, and it is not a suitable metric for whole-brain, cross-sectional analysis across individuals, but it can be used to make a node-wise, group average assessment analogous to that for weighted degree.

2.3.5 Tau Burden Assessment

Other than the atlas used to define regions of interest, all AV-1451 and PiB data acquisition and pre-processing steps were identical to those reported by Passamonti et al. (2017). Ligand preparation was carried out at the Wolfson Brain Imaging Centre, University of Cambridge, with high radiochemical purity. $[^{18}\text{F}]AV-1451$ was produced with a specific activity of 216±60 GBq/μmol at the end of synthesis (60min); $[^{11}\text{C}]$ PiB specific activity was >150GBq/μmol. PET imaging was performed on a GE Advance PET scanner with a 15min 68Ge transmission scan used for attenuation correction acquiring 58 frames of increasing duration. PET scanning was performed in 3D mode (63 image planes; 15.2cm axial field of view; 5.6mm transaxial resolution and 2.4mm slice interval) 80-100 minutes after a 9.0 to 11.0 mCi bolus injection in frames of 4x5 minutes. Each $[^{11}\text{C}]$ PiB scan was acquired using an 8.5 to 15 mCi bolus injection immediately followed by a 60-minute dynamic acquisition in 69 frames (12x15 seconds, 57x60 seconds). Scans were reconstructed on the GE Advance scanner using the PROMIS 3D filtered back projection algorithm, correcting for randoms, dead time, normalisation, sensitivity and scatter attenuation.

The cerebellar grey matter was used as a reference tissue to express the distribution volume ratio (DVR) for the $[^{11}\text{C}]$ PiB PET data for each MCI participant. PiB scans were classified as positive if the average SUVR values across the cortex were more than 1.5 times that of the cerebellar ROIs. Seven MCI participants met this criterion and were included in the study.

The non-displaceable binding potential of $[^{18}\text{F}]AV-1451$ was assessed at each ROI in the sub-parcellated atlas after rigid registration of each subject's dynamic PET image series to their T1-weighted MRI scan. This was normalised against the superior cerebellum, a reference region considered to have no tau pathology in either PSP or Alzheimer’s disease (Williams et al., 2007, Schöll et al., 2016, Passamonti et al., 2017).
These data were corrected for white matter and CSF partial volumes by calculating the ordinary least squared solution for the BP map voxel-wise in each region for grey plus white matter segments, each smoothed to PET resolution.

2.3.6 Statistical Approach

A stereotyped statistical approach was taken to the analysis of graph metrics of theoretical interest. All statistical analyses were performed in Matlab 2015b (The Mathworks Inc., 2015), with the exception of moderation analysis, which was performed in R (R Core Team, 2016).

A group-averaged analysis assessed hypothesis 1 by examining the node-wise relationships between $[^{18}\text{F}]\text{AV-1451}$ binding potential and weighted degree, weighted participation coefficient, and clustering coefficient. These analyses were performed across the whole brain and within ten intrinsic connectivity networks. To avoid circularity, where the regions of interest are defined by the test data, the intrinsic connectivity networks were independently derived from a publicly available dataset (Smith et al., 2009). A binary mask was constructed for each resting state network by thresholding at $Z \geq 2.6$ ($p<0.005$ uncorrected). Nodes were defined as belonging to a network if their centre of mass was within 5mm of any positive voxel.

A between-subject analysis assessed hypotheses 2 and 3. This was undertaken at a variety of spatial scales. I first looked across the whole brain. For each individual I first calculated a measure of disease-related tau burden. For participants with AD, in whom tau deposition increases in both magnitude and distribution as disease progresses (Braak and Braak, 1995), tau burden was calculated as average $[^{18}\text{F}]\text{AV-1451}$ binding potential across the whole brain. In PSP, tau deposition remains confined to brainstem and deep nuclei even in late disease (Williams et al., 2007); tau burden was therefore calculated as average $[^{18}\text{F}]\text{AV-1451}$ binding potential across the left and right thalamus, caudate, putamen, pallidum, accumbens, ventral diencephalon, midbrain, pons and medulla.. Both methods were separately assessed for controls. These measures of individual tau burden were then correlated with whole-brain averaged graph metrics to assess the relationship between the metric in question and disease.
burden in each group separately. Moderation analysis was performed for each metric to assess whether the relationships differed between AD and PSP. It has recently been reported that patients with early onset AD have a higher tau burden than those with later onset, especially in later Braak stage regions (Schöll et al., 2017). Non-significant trend relationships between overall tau burden and age were observed in our cohort (AD $r = -0.27$, PSP $r = -0.02$, Control $r = -0.38$; Figure 2-2). Cross-sectional analyses were therefore performed twice, with and without partialling out the effect of age on tau burden.

![Figure 2-2](image)

The relationship between tau burden and age for each group.

Both methods of calculating tau burden are shown for the control group. As previously reported, there was a negative trend in AD, although this was not statistically significant. No such relationship was demonstrated for PSP. Controls also displayed a negative relationship, although this was across a much smaller range of AV-1451 binding.
Regionality of the demonstrated effects was assessed within groups by correlating tau burden with the single subject graph metric values at each node; the gradient of the best fit linear regression within each group was the outcome measure. The nodal gradients in cerebellar regions were discarded (as this was the reference region for PET imaging), and the remaining regional maps were then collapsed into a vector. This was correlated with a matching vector of local tau burden at each node, calculated as the group-averaged increase in $[^{18}\text{F}]$AV-1451 binding potential. This resulted in a measure of the relationship between the distribution of disease-related change in each graph metric and regional deposition of tau.

2.4 Results

2.4.1 Resting state connectivity differences exist between groups

![Connectivity matrices](image)

**Figure 2-3** Connectivity matrices.
Upper: Association matrices derived from resting state (task-free) fMRI. Lower: Pairwise subtractions of the Fisher-z transformed association matrices to illustrate the large scale within and between-region disease related changes in functional connectivity.
As a prelude to the detailed analysis of network features, I first confirmed that our groups differed in their resting state connectivity, by testing for differences between the group averaged association matrices (Figure 2-3, upper). Jennrich tests confirmed highly significant statistical differences between controls and Alzheimer’s disease ($\chi^2 = 1.15 \times 10^8$, $n_1 = 12$, $n_2 = 17$, $P < 0.0001$); controls and PSP ($\chi^2 = 975547$, $n_1 = 12$, $n_2 = 17$, $P < 0.0001$); and Alzheimer’s disease and PSP ($\chi^2 = 733030$, $n_1 = 17$, $n_2 = 17$, $P < 0.0001$).

The nature of these differences would be opaque without abstraction techniques. Pairwise subtraction of the Fisher transformed association matrices (Figure 2-3, lower) revealed the presence of structure within the data, with both within-lobe and between-lobe group differences. While one can make general observations from these raw data, such as a prominent reduction in cerebello-cerebellar and brainstem-cerebellar connectivity in PSP, and a general reduction in between-region connectivity in Alzheimer’s disease, the insights would be limited. Instead, graph theoretic measures allow one to examine which specific properties of the network underpin disease-related differences. For this approach each region in the brain becomes a node in a graph, which is functionally connected to all other nodes with a finite strength that is given by their pairwise correlation over time. Each individual’s graph is then thresholded so that the highest ‘x’ percentage of connections survives. Effects are sought that are consistently present at a variety of network thresholds. Here I examine network thresholds from $1 < x < 10\%$, representing a range of graphs from sparse to dense. Very sparse graphs contain less information and can miss important relationships. Conversely, very dense graphs are more subject to noise and, when binarised, begin to provide less meaningful information. Therefore, in what follows, I present the primary statistical analyses at an intermediate density of 6\%, with statistical detail given for this density. To assess the robustness of effects to different thresholding decisions I note separately the range of network densities exhibiting statistical significance.

### 2.4.2 Findings in AD but not PSP are consistent with trans-neuronal spread of tau

I hypothesised that trans-neuronal spread of tau would manifest as a positive relationship between $[^{18}\text{F}]\text{AV-1451}$ and weighted degree. As I have parcellated the brain into nodes of equal size, weighted degree is a measure of the volume of cortex to
which a node is connected, and the strength of these connections. In Alzheimer’s
disease, a strong positive correlation was observed at all network thresholds, such that
the most strongly connected nodes had higher $[^{18}\text{F}]\text{AV-1451 BP}_{\text{ND}}$ (Pearson’s $r = 0.48$,
$P < 0.0001$, Spearman’s rho = 0.48, $P < 0.0001$) (Figure 2-4 A). A consistent
relationship was not present in PSP (Pearson’s $r = -0.09$, Spearman’s rho = 0.12)
(Figure 2-4 B) or controls (Pearson’s $r = 0.03$, Spearman’s rho = 0.11) (Figure 2-4 C).
These patterns were present at all examined network density thresholds.
Figure 2–4 Weighted degree.
A–C: Group-averaged connection strength at each node, quantified by weighted degree, plotted against $[^{18}\text{F}]\text{AV-1451}$ binding potential at that node. A statistically significant linear relationship was demonstrated only in AD (Pearson’s $r = 0.48$, $p < 0.0001$, Spearman’s rho $= 0.48$, $p < 0.0001$), and the corresponding regression line is plotted for this group. D: Between-subjects analysis of the relationship between global tau burden and each weighted degree at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden – for simplicity only their whole-brain average points are illustrated here. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (AD and PSP) was statistically significant. E: The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups. F: Average $[^{18}\text{F}]\text{AV-1451}$ binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of
each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. G: The local tau burden-related change in weighted degree is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. H: Average raw values of each graph metric within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.

It is important to acknowledge that the degree of a node as defined by correlated activity is impacted by the size of the connectivity network to which it belongs (Power et al., 2013). Therefore, one concern might be that correlations between $[^{18}\text{F}]$AV-1451 and degree are driven by the coincidence that AD happens to affect the default mode network (DMN), which happens to be a large intrinsic connectivity network. To exclude this possibility, I undertook two additional analyses. Firstly, I re-examined $[^{18}\text{F}]$AV-1451 and degree when nodes belonging to the DMN were excluded; this did not abolish the strong relationship (Pearson’s $r = 0.43$, $P < 0.0001$). Secondly, I examined the relationship within each intrinsic connectivity network (Figure 2-5). A strong positive relationship between $[^{18}\text{F}]$AV-1451 and degree was seen in nine out of the ten networks examined (Pearson’s $r >= 0.4$, $P < 0.0001$).
Figure 2-5 Alzheimer’s disease comparison of the three graph metrics representing the three principal hypotheses of hub vulnerability. Broken down by intrinsic connectivity network defined from (Smith et al., 2009). The group-averaged graph metric at each node within a network is plotted against \([^{18}F]AV-1451\) binding potential at that node. The Pearson correlation coefficient is noted in each case. Only weighted degree demonstrated a consistent relationship across all networks in keeping with its related hypothesis.
In contrast, no strong positive correlations were observed in any network in PSP (Figure 2-6).

Figure 2-6 PSP comparison of the three graph metrics representing the three principal hypotheses of hub vulnerability.

*Broken down by intrinsic connectivity network defined from (Smith et al., 2009). The group-averaged graph metric at each node within a network is plotted against $[^{18}\text{F}]$AV-1451 binding potential at that node. The Pearson correlation coefficient is noted in each case.*

Finally, in each group I examined the relationship between each node’s unthresholded connectivity strength and $[^{18}\text{F}]$AV-1451 BP$_{ND}$ (Figure 2-7). This is a local measure of the strength of functional connectivity that does not rely on thresholding the graph and therefore takes into account both strong and weak connections. The pattern of results observed in weighted degree was replicated; in Alzheimer’s disease, a positive correlation was observed ($r = 0.28, P < 0.0001$), but no significant relationship was observed in PSP ($r = -0.05, P = 0.22$) or controls ($r = 0.04, P = 0.31$).
Figure 2-7 Unthresholded nodal connectivity strength.
A: Group average for Alzheimer’s disease. A statistically significant positive relationship was observed ($r = 0.28$, $p < 0.0001$). B: The disease-related change in nodal connectivity strength at each node in the Alzheimer’s disease group. A statistically significant negative relationship was observed ($r = -0.34$, $p < 0.0001$). Therefore nodes that were more strongly connected accrued more tau, but the consequence of tau accrual was that those same nodes then lost connectivity. C: Group average for PSP. No statistical relationship was observed ($r = -0.05$, $p = 0.22$). D: Group average for controls. No statistical relationship was observed ($r = 0.04$, $p = 0.31$).
2.4.3 Findings in PSP but not AD are consistent with hub vulnerability due to metabolic demand and lack of trophic support.

I hypothesised that tau accumulation due to metabolic demand would manifest as a positive relationship between [18F]AV-1451 and weighted participation coefficient. This was not observed in AD; in fact there was a weak negative correlation (Pearson’s r = −0.18) (Figure 2-8 A).

In PSP, however, those nodes that displayed elevated [18F]AV-1451 were those that had the highest participation coefficient (Figure 2-8 B).
Figure 2-8 Weighted participation coefficient.
A-C: Group-averaged weighted participation coefficient at each node, plotted against [18F]AV-1451 binding potential at that node. A negative relationship was observed in Alzheimer’s disease, in violation of the metabolic demand hypothesis. D: Between-subjects analysis of the relationship between global tau burden and participation coefficient at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (AD and PSP) was statistically significant. E: The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups. F: Average [18F]AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. G: The local tau burden-related change in participation coefficient is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range.
of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. H: Average raw values for participation coefficient within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.

Similarly, I hypothesised that tau accumulation due to a lack of trophic support would manifest as a negative relationship between [18F]AV-1451 and clustering coefficient. In AD, the opposite relationship was observed; strongly clustered nodes were more likely to display elevated [18F]AV-1451 binding (Pearson’s r = 0.33) (Figure 2-9 A).

Again, by contrast, in PSP those nodes that displayed elevated [18F]AV-1451 were those that had the lowest clustering coefficient (Figure 2-9 B).
Figure 2-9 Clustering coefficient.
A-C: Group-averaged clustering coefficient at each node, plotted against [18F]AV-1451 binding potential at that node. A positive relationship was observed in Alzheimer’s disease, in violation of the trophic support hypothesis. D: Between-subjects analysis of the relationship between global tau burden and clustering coefficient at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (AD and PSP) was statistically significant. E: The magnitude of disease-related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups. F: Average [18F]AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. G: The local tau burden-related change in clustering coefficient is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease
burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. H: Average raw values for clustering coefficient within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.

2.4.4 Alzheimer’s disease and PSP have opposite effects on the strength of cortical functional connectivity

I assessed the impact of the presence of tau on connection strength at a variety of spatial scales. Firstly, I averaged weighted degree across the whole brain, resulting in a single measure for each individual. As the overall number of connections in each individual’s graph was thresholded at an identical network density, this measure represented the average strength of the strongest X% of connections. To assess global disease burden in Alzheimer’s disease I averaged $[^{18}\text{F}]\text{AV}-1451\text{ BP}_{\text{ND}}$ across the whole brain. In PSP, I averaged $[^{18}\text{F}]\text{AV}-1451\text{ BP}_{\text{ND}}$ across midbrain and basal ganglia, reflecting the more focal distribution of disease and our hypotheses about dissociated functional effects of cortical and subcortical tau (Passamonti et al., 2017). Confirmatory tests demonstrated that the pattern of the moderation analyses below were unchanged if whole-brain average $[^{18}\text{F}]\text{AV}-1451\text{ BP}_{\text{ND}}$ was used to assess tau burden in PSP. Both methods of assessing tau burden were independently assessed in the control group, with no significant effects demonstrated in either case.

In AD, I found a negative correlation between average connection strength and tau burden (Pearson’s $r = -0.58$, $p = 0.015$) (Figure 2-4 D). In PSP this relationship was reversed, with average connection strength increasing in line with tau burden ($r = 0.65$, $p = 0.004$). No relationship was observed in controls ($r = -0.10$, $p = 0.75$). Moderation analysis confirmed a dissociated relationship between tau and connection strength in Alzheimer’s disease and PSP at all network densities ($\Delta r^2 = 0.29$, $F(1,30) = 19.0$, $p = 0.0001$; after partialling out age (Figure 2-10 A) $\Delta r^2 = 0.28$, $F(1,30) = 17.9$, $p = 0.0002$). Therefore, while in Alzheimer’s disease the presence of tau pathology causes the strongest functional connections to weaken, in PSP, the presence of tau had the opposite effect, i.e., to strengthen functional connections.
Secondly, I assessed the distribution of change by repeating the correlation of disease burden against weighted degree at every individual node. I hypothesised that this effect would be greatest in those regions that display the strongest functional connectivity in the healthy brain, and which I have demonstrated to accrue most tau in Alzheimer’s disease. The gradient of this nodewise relationship reflects a measure of local change in weighted degree with disease burden (Figure 2-4 E). Firstly, I examined whether the whole-brain average relationship could be replicated in these individual gradients, by performing sign tests. For Alzheimer’s disease, a negative relationship was confirmed (Z = −13.0, P < 0.0001); for PSP there was a positive relationship (Z = 14.8, P < 0.0001); while for controls no relationship was demonstrated using either whole brain (Z = −1.1, P = 0.27) or deep brain (Z = 1.1, P = 0.27) tau burden.

Next, I assessed whether the functional connectivity change at each node related to local tau burden, by correlating the gradient of the disease-related change in weighted degree with the disease-associated increase in [18F]AV-1451 binding potential at each node (i.e. correlating each column of Figure 2-4 F with the corresponding column of Figure 2-4 G). A negative correlation between these measures was demonstrated in Alzheimer’s disease (Pearson’s r = −0.30, P < 0.0001, Spearman’s rho = −0.24, P < 0.0001). This relationship was absent in PSP (Pearson’s r = −0.07, P = 0.11, Spearman’s rho = −0.00, P = 0.98) and in controls (Pearson’s r = −0.01, P = 0.75, Spearman’s rho = −0.01, P = 0.80). As well as being present at all examined network densities, the negative correlation between disease-related change in functional connectivity strength and tau burden in Alzheimer’s disease was replicated in nodal connectivity strength, the equivalent unthresholded measure (r = −0.34, P<0.0001, Figure 2-7 B).

Finally, I assessed whether the functional connectivity change at each node related to the strength of its connections in the healthy control brain (i.e. correlating each patient panel of Figure 2-4 G with the control panel in Figure 2-4 H). As would be expected from the propensity of highly connected nodes to accrue tau, a negative relationship was demonstrated in Alzheimer’s disease (Pearson’s r = −0.23, P < 0.0001, Spearman’s rho = −0.24, P < 0.0001). Importantly, however, this relationship explained less variance than AV binding, with which I have demonstrated it to be correlated. In PSP,
a positive relationship was demonstrated (Pearson’s $r = 0.27$, $P < 0.0001$, Spearman’s rho $= 0.26$, $P < 0.0001$).

In summary, nodes that are constitutionally more strongly connected to a larger volume of cortex are more likely to accrue tau pathology in Alzheimer’s disease but not PSP. This relationship is independent of the connectivity network to which a node belongs. Once present, the tau pathology appears to cause local functional connectivity strength to fall. By contrast, in PSP tau selectively accumulates in midbrain and deep nuclei, which has the consequence of increasing the strength of cortico-cortical functional connectivity, especially in those nodes that are constitutionally highly connected.
Figure 2-10 Between-subjects analysis of the relationship between global tau burden and each graph metric at a network density of 6%, with the effect of age on tau burden partialled out. A Relative tau burden of 0 is the age-expected tau burden within each disease group (i.e. lying on the group trend line in Figure 2-2). Individuals with lower-than-average tau burden within their disease group have negative Relative tau burdens, while those with higher-than-average tau burden have positive Relative tau burdens. Moderation...
analysis for a differential relationship between graph metric and tau burden in the two disease groups (AD and PSP) was statistically significant for all metrics.

2.4.5 Reorganisation of cortical functional connectivity reflects cortical vs subcortical pathology

I hypothesised that the presence of subcortical pathology in PSP might be causing an increase in weighted degree by necessitating an increase in the relative strength of short-range cortico-cortical connections as longer range connections are disrupted. This view is supported by the observation that a consequence of increasing tau burden in PSP is a marked reduction in participation coefficient (Figure 2-8 D). This relationship was not observed in AD, where increasing tau burden caused increasing participation. Moderation analysis confirmed this differential relationship at all network densities from 1-10% (Δr² = 0.21, F(1,30) = 7.6, p = 0.002; after partialling out age (Figure 2-10 B) Δr² = 0.17, F(1,30) = 8.3, p = 0.007).

I further examined this hypothesis by examining the effect of tau on other measures of graph structure in each individual, dissociated from variation in degree by binarisation after thresholding and normalisation against 1000 random graphs with the same number of connections at each node.

The clustering coefficient quantifies how many of a node’s neighbours are neighbours of each other. All individuals in all groups displayed clustering to at least 2.8x that of random graphs of the same degree (Figure 2-9 D), consistent with a small-world connectivity distribution. While in AD a non-significant trend was observed towards reduced clustering with increasing tau burden, in line with previous reports (Stam et al., 2006, Sanz-Arigita et al., 2010), in PSP the opposite relationship was observed. Moderation analysis trended towards significance at the density of primary interest (Δr² = 0.08, F(1,30) = 3.8, p = 0.06; after partialling out age (Figure 2-10 C) Δr² = 0.07, F(1,30) = 2.5, p = 0.09), and was significant from 7-10% density. The betweenness centrality of a node is a measure of the number of shortest paths between any other two nodes that pass through it. If the presence of subcortical pathology in PSP means that long-range information transfer must occur through a trans-cortical route, average betweenness centrality should increase. Conversely, predominantly cortical pathology in AD might increase reliance on cortico-subcortical connections,
reducing average betweenness centrality. This prediction was verified (Figure 2-11 A), with a differential relationship between tau burden and betweenness centrality confirmed by moderation analysis at all network densities from 1-10% ($\Delta r^2 = 0.21$, $F(1,30) = 9.3$, $p = 0.005$; after partialling out age (Figure 2-10 D) $\Delta r^2 = 0.20$, $F(1,30) = 8.9$, $p = 0.006$).
Figure 2-11 Betweenness centrality.

A: Between-subjects analysis of the relationship between global tau burden and betweenness centrality at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (AD and PSP) was statistically significant. B: The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups (supplementary table 2). C: Average [18F]AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. D: The local tau burden-related change in betweenness centrality is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. E: Average raw values for betweenness centrality within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.
Long-range information transfer by a cortico-cortical route implies an inefficient, indirect process, perhaps accounting for the cognitive slowing characteristic of neurodegenerative disorders with predominant sub-cortical pathology. I tested for this by examining closeness centrality, which is the inverse of the path length between a node and all other nodes in the graph. As would be predicted by this account, increasing disease burden resulted in a higher average path length (lower closeness centrality) in PSP but a lower average path length in Alzheimer’s disease (higher closeness centrality) at network densities from 2-10% (Figure 2-12 A, moderation Δr² = 0.20, F(1,30) = 9.1, p = 0.005; after partialling out age (Figure 2-10 E) Δr² = 0.16, F(1,30) = 8.0, p = 0.008). Averaged across the whole brain, this measure is equivalent to the global efficiency of the graph, and indeed the same statistical results were obtained from a moderation analysis of that metric; increasing disease burden resulted in reduced global efficiency in PSP.
Figure 2-12 Closeness centrality (1/path length from each node to all other nodes).
A: Between-subjects analysis of the relationship between global tau burden and closeness centrality at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (AD and PSP) was statistically significant. B: The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups (supplementary table 2). C: Average [18F]AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. D: The local tau burden-related change in closeness centrality is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. E: Average raw values for closeness centrality within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.
While tau-mediated re-organisation of connectivity in PSP results in slower and less efficient long-range information transfer, in Alzheimer’s disease one might suppose it to be a beneficial consequence of the loss of overall connection strength as those connections that remain are more globally efficient. However, I hypothesised this would be at the cost of local efficiency, which is a measure of the number of strong connections between neighbouring nodes and the robustness of local networks to disruption. Indeed this was found to be the case at network densities from 4-10% (Figure 2-13 A, $\Delta r^2 = 0.15$, $F(1,30) = 6.5$, moderation $p = 0.016$; after partialling out age (Figure 2-10 F) $\Delta r^2 = 0.12$, $F(1,30) = 5.5$, $p = 0.026$).
Figure 2-13 Local efficiency.
A: Between-subjects analysis of the relationship between global tau burden and local efficiency at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (AD and PSP) was statistically significant. B: The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups (supplementary table 2). C: Average [18F]AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. D: The local tau burden-related change in local efficiency is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. E: Average raw values for local efficiency within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.
Finally, I assessed the functional influence of each node on the other nodes in the network by examining the eigenvector centrality. Again, an opposite effect of disease burden was observed in AD and PSP at network densities from 3-10% (Figure 2-14 A, \( \Delta r^2 = 0.22, F(1,30) = 8.5, p = 0.007 \); after partialling out age (Figure 2-10 G) \( \Delta r^2 = 0.20, F(1,30) = 8.3, p = 0.007 \)) such that, on average, as PSP progressed each node had less functional influence on every other node in the graph.
Figure 2-14 Eigenvector centrality.
A: Between-subjects analysis of the relationship between global tau burden and eigenvector efficiency at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (AD and PSP) was statistically significant. B: The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups (supplementary table 2). C: Average [18F]AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. D: The local tau burden-related change in eigenvector efficiency is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. E: Average raw values for eigenvector efficiency within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.
2.4.6 Local functional connectivity reorganisation is related to local $[^{18}\text{F}]$AV-1451 binding potential in Alzheimer’s disease

To understand the differential reorganisation of functional connectivity in Alzheimer’s disease and PSP, I replicated the approach taken for weighted degree; correlating the gradient of the disease-related change at every individual node with local change in $[^{18}\text{F}]$AV-1451 binding potential. For all metrics, the differential group effects demonstrated by moderation analysis could be replicated by statistical significance of opposite directionality in sign tests performed on the individual gradients at each node (Table 2-3, Figure 2-4, Figure 2-8, Figure 2-9, Figure 2-11, Figure 2-12, Figure 2-13, Figure 2-14). In all cases, controls showed either no relationship or a weaker effect in the same direction as patients with Alzheimer’s disease.

<table>
<thead>
<tr>
<th>Group</th>
<th>Betweenness Centrality</th>
<th>Closeness Centrality</th>
<th>Local Efficiency</th>
<th>Eigenvector Centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>Z = - 5.4, p&lt;0.0001</td>
<td>Z = 19.8, p&lt;0.0001</td>
<td>Z = - 11.2, p&lt;0.0001</td>
<td>Z = 6.3, p&lt;0.0001</td>
</tr>
<tr>
<td>PSP</td>
<td>Z = 5.0, p&lt;0.0001</td>
<td>Z = - 17.9, p&lt;0.0001</td>
<td>Z = 9.3, p&lt;0.0001</td>
<td>Z = - 10.8, p&lt;0.0001</td>
</tr>
<tr>
<td>Control (tau burden averaged across brain as for AD)</td>
<td>Z = - 0.45, p=0.65</td>
<td>Z = 5.8, p&lt;0.0001</td>
<td>Z = - 6.9, p&lt;0.0001</td>
<td>Z = 3.9, p&lt;0.0001</td>
</tr>
<tr>
<td>Control (tau burden assessed in brainstem and deep nuclei as for PSP)</td>
<td>Z = 0, p=1</td>
<td>Z = 9.7, p&lt;0.0001</td>
<td>Z = - 4.2, p&lt;0.0001</td>
<td>Z = 3.6, p=0.0002</td>
</tr>
</tbody>
</table>

Table 2-3 Sign tests
Performed on the best fit linear gradients at each of 598 nodes, comparing tau burden and graph metric across individuals. Analysis was performed twice in controls, independently assessing the measures of tau burden used in AD and PSP.

The reorganisation of graph metrics followed two distinct patterns. Closeness centrality (Figure 2-12) displayed a global effect, with most nodes increasing in Alzheimer’s disease and decreasing in PSP. This was not strongly related to local $[^{18}\text{F}]$AV-1451 binding potential in AD (Pearson’s $r = 0.05$, $p = 0.25$, Spearman’s rho = 0.06, $p = 0.18$) or PSP (Pearson’s $r = 0.08$, $p = 0.06$, Spearman’s rho = 0.11, $p = 0.01$). Similarly,
participation coefficient displayed global changes in both diseases (Figure 2-8) (AD Pearson’s $r = 0.09$, Spearman’s rho = 0.07, PSP Pearson’s $r = 0.08$, Spearman’s rho = 0.06).

By contrast, eigenvector centrality (Figure 2-14) was more strongly related to local $[^{18}\text{F}]$AV-1451 binding potential in Alzheimer’s disease (Pearson’s $r = -0.28$, $P < 0.0001$, Spearman’s rho = $-0.25$, $P < 0.0001$). Strikingly, the positive relationship I demonstrated across the whole brain masked opposing regional effects. As global tau burden increased, the functional influence of frontal regions on all other regions increased, while that of occipital regions decreased. In PSP no consistent relationship with AV was observed (Pearson’s $r = -0.10$, $p = 0.02$, Spearman’s rho = $-0.04$, $p = 0.38$), with almost all regions having less functional influence on all other regions as tau burden increased.

Clustering coefficient (Figure 2-9), local efficiency (Figure 2-13) and betweenness centrality (Figure 2-11) displayed an intermediate degree of regional specificity, being weakly but significantly correlated with change in $[^{18}\text{F}]$AV-1451 binding potential in AD (clustering coefficient Pearson’s $r = -0.12$, $p<0.0001$, Spearman’s rho = $-0.14$, $p<0.0001$, local efficiency Pearson’s $r = -0.16$, $p<0.0001$, Spearman’s rho = $-0.19$, $p<0.0001$, betweenness centrality Pearson’s $r = 0.19$, $p<0.0001$, Spearman’s rho = $0.19$, $p<0.0001$) and PSP (clustering coefficient Pearson’s $r = -0.14$, $p<0.0001$, Spearman’s rho = $-0.12$, $p<0.0001$, local efficiency Pearson’s $r = -0.14$, $p = 0.001$, Spearman’s rho = $-0.11$, $p = 0.02$, betweenness centrality Pearson’s $r = 0.17$, $P < 0.0001$, Spearman’s rho = $0.08$, $P = 0.08$).

2.5 Discussion

I have demonstrated that in Alzheimer’s disease a strong relationship exists between the propensity of a node to display elevated $[^{18}\text{F}]$AV-1451 binding and the volume of cortex to which it is strongly connected (Figure 2-4 A). Further, I have demonstrated that this effect exists both within and between intrinsic connectivity networks (Figure 2-5). This is consistent with the theory of trans-neuronal spread. The predictions of the competing hypotheses that highly active brain regions are vulnerable to tau accumulation due to a positive relationship with metabolic demand (Figure 2-8 A) or
negative relationship with clustering due to a lack of trophic support (Figure 2-9 A) were not supported by the data.

In contrast, in PSP, I have demonstrated the opposite findings. Those brain regions that accrue most tau display weak connectivity (Figure 2-4 B), but are predicted to have high metabolic demand (Figure 2-8 B) and a lack of trophic support (Figure 2-9 B).

Further, I have explored the consequences of tau accumulation with cross sectional analyses at a variety of spatial scales. In Alzheimer’s disease, I have demonstrated that with greater levels of tau pathology the strongest inter-nodal connections are weakened (Figure 2-4 G). This reorganisation of the brain network leads to more direct long-range connections passing through fewer nodes (Figure 2-12 A, Figure 2-11 A), at the cost of lower local efficiency (Figure 2-13 A).

In PSP, where tau accumulation is predominantly subcortical, I demonstrate the opposite re-organisation of the connectivity graph. With greater levels of tau in midbrain and deep nuclei (Figure 2-4 F), the strongest functional connections are strengthened; these are predominantly cortical (Figure 2-4 G, H). Information transfer therefore takes a less direct path (Figure 2-12 D), passing through a larger number of cortical nodes en-route (Figure 2-11 A), as deep structures can no longer sustain long range connectivity. This is accompanied by a decrease in participation coefficient (Figure 2-8 D) and an increase in clustering (Figure 2-9 D) as connectivity becomes increasingly modular.

2.5.1 Insights into the mechanisms of disease progression in humans

It has been proposed that the pathological mechanisms underlying Alzheimer’s disease begin in a single, vulnerable location and spread from cell to cell, rather than occurring independently in a large number of vulnerable cell populations (Guo and Lee, 2014, Goedert, 2015). The primary direct evidence for such propagation of tau comes from rodent studies. For example, the injection of brain extract from transgenic mice expressing mutant tau into mice expressing wild-type human tau caused wild-type tau to form filaments and spread to neighbouring brain regions (Clavaguera et al., 2009). Further, pathological tau from human brains causes disease in wild-type mice, in whom the pathological human tau species becomes self-propagating (Clavaguera et al., 2013).
This tau propagation is mediated by the presence and strength of synaptic connectivity rather than spatial proximity (Liu et al., 2012, Iba et al., 2013, Ahmed et al., 2014).

Associative studies of the healthy brain have demonstrated that large-scale, functionally connected neural networks strongly resemble the known patterns of atrophy in distinct neurodegenerative syndromes mediated by tau and TDP-43 (Zhou et al., 2012). Further human evidence comes from the observation that patterns of atrophy in the rare disease non-fluent variant primary progressive aphasia (nfvPPA) strongly correspond to structural and functional connectivity in the healthy speech production network (Mandelli et al., 2016). This is an important observation, because clinically and radiologically indistinguishable cases of nfvPPA can be caused either by tau or by the unrelated protein TDP-43, which has been demonstrated to propagate trans-neuronally (Braak et al., 2013).

Here, I go beyond these associative studies to measure tau burden and functional connectivity in the same individuals at both the whole-brain and regional level. Our observation that those brain areas that are more strongly functionally connected have accrued more tau pathology in Alzheimer’s disease (Figure 2-4 A), independent of which connectivity network they belong to (Figure 2-5), is consistent with trans-neuronal spread. I demonstrate that the presence of tau is not, in itself, inducing stronger regional connectivity by our cross sectional analysis of the AD group, in which I demonstrate that as cortical tau accumulates the overall functional connectivity of cortex falls (Figure 2-4 D), and this between-subjects effect is strongest in those brain regions with most tau accumulation (Figure 2-4 F, G). Crucially, I demonstrate that [¹⁸F]AV-1451 binding potential at each node is better than the connectivity of that node in the healthy brain at accounting for regional variance in connectivity change, arguing against the presence of tau being a secondary marker of neurodegeneration in vulnerable hubs. In other words it is not coincidence that AD tends to impact large networks; it is a predictable consequence of trans-neuronal spread of a disease-causing protein. Graph theoretic models of transmissible disease epidemics are in agreement that the likelihood of an individual becoming infected (and the dose of the infectious agent received) is directly proportional to its number of infected neighbours and their infectivity (Durrett, 2010). As our nodes represent brain regions of equal volume, the binary portion of degree represents a surrogate measure of the number of neurons to
which a brain region is connected, and the weighted portion of degree is a measure of
the strength of these connections. By the time Alzheimer’s Disease is sufficiently
advanced to cause the symptoms of mild cognitive impairment, tau is generally already
present to some degree throughout neocortex (Markesbery, 2010), and therefore
reaching a disease stage at which the number of neighbours more closely approximates
the number of infected neighbours, and the connection strength between infected
neighbours (here the weighted portion of degree) becomes a strong driver of infectivity.

Conversely, our analyses do not provide support for models of hub vulnerability due to
metabolic demand or lack of trophic support in Alzheimer’s disease. It is important to
acknowledge that this does not mean that these mechanisms are unimportant, but
rather that they are a downstream event of tau accumulation. In other words, while I
demonstrate that the propensity of a node to accrue tau is not related to metabolic
demand or trophic support, these factors might still contribute to determining the
vulnerability of brain regions to the presence of a given amount of tau. This hypothesis
could be addressed in future studies by relating the information content of tau ligand
binding to other measures of neurodegeneration such as longitudinal changes in grey
matter volume.

In PSP, I do find support for models of hub vulnerability due to metabolic demand and
lack of trophic support, but not for models of trans-neuronal spread. This is in line
with other recent studies, demonstrating that neurodegeneration in PSP is related to
local gene expression patterns (Rittman et al., 2016).

2.5.2 Is the difference between AD and PSP mediated by tau isoform or intrinsic
connectivity?

A striking feature of our results is that I demonstrate a strong relationship between
[18F]AV-1451 binding potential and the strength of functional connectivity in AD but
not PSP. There are at least two potential (and not mutually exclusive) explanations for
this dissociation.

The first possibility is that trans-neuronal tau propagation might occur more slowly in
PSP. Cellular models have demonstrated that the propensity of tau to propagate
intracellularly depends on its ability to form aggregates (Falcon et al., 2015). Once
present in a new cell, the ability of tau fibrils to induce the aggregation of constitutionally present tau depends on the conformation of the fibril structure (Nonaka et al., 2010, Fitzpatrick et al., 2017). It might be that the straight filaments of predominantly 4R tau that characterise PSP are simply less able to propagate or need to be present in higher concentrations before they can induce a chain reaction of local tau aggregation (Guo et al., 2016). This view is supported by our demonstration that tau is restricted to brainstem and deep nuclei in PSP despite these nodes being highly promiscuous between networks, displaying a high participation coefficient (Figure 2-8 B) and low clustering (Figure 2-9 B).

Secondly, it is possible that tau propagation does occur, but that it is limited in range to a subcortical intrinsic connectivity network (Raj et al., 2012) (perhaps network 5 in (Laird et al., 2011)). Such a subcortical network may be poorly visualised by multi-echo fMRI, and is less frequently observed at rest than the default mode network (Greicius et al., 2003), which has been implicated in Alzheimer’s disease (Greicius et al., 2004). Together, therefore, our findings of weak functional connectivity in deep nuclei and brainstem nodes (Figure 2-4 B) might mask meaningful functional connectivity within and between these regions, accounting for the restricted and stereotyped pattern of tau accumulation in early PSP (Ellison et al., 2012) and cortical escape in advanced PSP (Schofield et al., 2012a).

2.5.3 Cross-sectional data reveal patterns less visible at the group level

Computational modelling of connectivity-dependent cell death predicts a dissociation between changes in functional connectivity in early and late Alzheimer’s disease (de Haan et al., 2012). It has been proposed that, in early disease, hubs compensate for declines in structural connectivity by increasing their firing rate, manifesting as stronger functional connectivity (Maestú et al., 2015). As disease progresses, this mechanism breaks down as neural damage prevents the maintenance of this metabolically demanding compensation (Jones et al., 2015). This dissociation is thought to underlie some of the seemingly inconsistent findings in the analysis of graph properties in neurodegenerative disease. By using a cross-sectional approach across a range of disease severity, I demonstrate relationships consistent with this hypothesis. In Alzheimer’s disease, weighted degree (Figure 2-4 D) and betweenness centrality (Figure
2-11 A) consistently fall as tau burden increases. However, in early disease where tau burden is low (our sample includes a range of severity, including PiB positive mild cognitive impairment), the regression line for these metrics is above the control average. Similarly, examining the regional changes related to disease burden reveals striking patterns that are obscured at the group level. This underlines the more general principle that one should be cautious in interpreting a main effect of group in the presence of an interaction, or correlation with severity.

2.5.4 Re-organisation of brain networks

Our examination of two distinct tau-mediated neurodegenerative pathologies with different distributions of pathology has enabled us to distinguish their consequences. I demonstrate strongly opposing effects in a range of metrics resulting from the presence of predominantly cortical (AD) or subcortical (PSP) tau. The direction of these effects in Alzheimer’s disease is consistent with the previously recognised impact of neurodegeneration, leading to increasingly random cortical connectivity (Sanz-Arigita et al., 2010, Stam, 2014) and a reduction in small-world properties (Stam et al., 2006). Eigenvector centrality showed particularly strong regional effects in Alzheimer’s disease, with a negative correlation observed between disease-related changes in this metric and local tau burden at each node – i.e. those brain regions that displayed less tau pathology had greater functional influence on other brain regions.

In contrast, cortico-subcortical functional connectivity is preferentially impaired in PSP-Richardson’s syndrome, resulting in cerebral information transfer taking a less direct path through a larger number of cortical nodes, reducing closeness centrality and eigenvector centrality, but increasing cortical degree, betweenness centrality and local efficiency. This results in an excessively modular connectivity arrangement, with decreasing participation coefficient and increasing clustering. These findings tie together classical observations of the ‘subcortical dementia’ phenotype of PSP (Albert et al., 1974) with more modern observations that behavioural change and cognitive impairment in PSP correlates with frontal cortical hypometabolism (D’Antona et al., 1985, Foster et al., 1988) and atrophy (Cordato et al., 2002, Cordato et al., 2005). I propose that increases in cortical functional connectivity can compensate for subcortical tau burden in PSP. Increasingly indirect information transfer accounts for
the cognitive slowing that is the hallmark of the ‘subcortical dementias’, but performance on untimed tests is preserved until cortical regions become atrophic in late disease.

2.5.5 Study limitations

The main limitation of our analysis is that it is cross-sectional, and I use $^{18}$FAV-1451 binding as a surrogate marker of tau burden. By making observations about the relationship between tau burden and functional connectivity in this way, I assume a uniformity of effect within our disease groups. As novel PET ligands such as $^{18}$FAV-1451 gain maturity, longitudinal assessment of tau burden and functional connectivity in the same individuals will be an important and powerful validation of our results. Definite evidence of the causal relationship between tau and connectivity with require the combination of longitudinal assessment and interventional studies targeting tau pathology. It should also be noted that $^{18}$FAV-1451 binding identifies predominantly aggregated tau in tangles, and does not directly measure oligomeric tau, which may be more toxic to the cell and synaptic plasticity, nor extracellular forms of tau that may mediate spread of pathology. The molecular binding target of $^{18}$FAV-1451 in non-AD tauopathies is disputed; elevated ‘off-target’ binding has been demonstrated in the basal ganglia of healthy controls (Johnson et al., 2016), albeit to a lesser degree than that observed in PSP (Figure 2-4 F, (Passamonti et al., 2017)), and in TDP-43 associated disorders without evident tau pathology (Bevan-Jones et al., 2017a). Nonetheless, $^{18}$FAV-1451 is able to recapitulate the distribution of post-mortem neuropathology in these disorders, making it appropriate for use here (Smith et al., 2016).

Our analysis is focussed towards cortico-cortical functional connectivity. In particular, multi-echo MRI might have a poor signal to noise ratio in deep brain structures, although the main advantage of using this sequence is that it enables robust de-noising of movement related artefacts pipeline (Kundu et al., 2012, Kundu et al., 2013). This is critical in clinical populations, in which fMRI data may differentially suffer from quality degradation due to head movements.

Finally, by examining proportionately thresholded graphs with 1-10% density, our analysis focusses on the strongest inter-regional functional connections. However, it is
possible that I am missing additional effects of neurodegeneration on weak or medium-strength connections. Tract-tracing studies indicate that there are weak anatomical connections, equivalent to a few axons, between some cortical areas (Ypma and Bullmore, 2016). Such weak links may have functional importance in complex networks (Granovetter, 1983). However, weak connections are difficult to evaluate with functional MRI, as it is not possible to disentangle them from correlation arising from signal noise. Future evaluation of these weaker connections with in vivo tractography, neuropathology or novel methods might reveal additional effects not evident in our dataset. In the interim, the thresholding procedure exhibits several advantages; by retaining only the most strongly correlated edges one is less likely to include false positive correlations and topologically random edges. It also allows the computationally intense process of normalisation of metrics against random graphs of equal density. The consistency between the results using thresholded nodal weighted degree and unthresholded nodal connectivity strength provides reassurance in the choice of thresholding of connections.

2.6 Conclusion

This study reveals the differential relationship between tau burden and functional connectivity in two distinct human neurodegenerative tauopathies. Our results enable us to disentangle the causes of tau accumulation from their consequences. They have wide-ranging implications, from the validation of models of tau trafficking in humans to corroborating computational models of hub compensation in Alzheimer’s disease, while accounting for the contrasting cognitive phenotype of these two conditions. These insights into the relationship between tau burden and brain connectivity changes will inform translational models and clinical trials of disease-modifying therapies.
2.7 Chapter Summary:

Alzheimer’s Disease (AD) and Progressive Supranuclear Palsy (PSP) represent neurodegenerative tauopathies with predominantly cortical vs subcortical disease burden. In AD, neuropathology and atrophy preferentially affect ‘hub’ brain regions that are densely connected. It was unclear whether hubs are differentially affected by neurodegeneration because they are more likely to receive pathological proteins that propagate trans-neuronally, in a prion-like manner, or whether they are selectively vulnerable due to a lack of local trophic factors, higher metabolic demands, or differential gene expression.

I assessed the relationship between tau burden and brain functional connectivity, by combining in vivo PET imaging using the ligand AV-1451, and graph theoretic measures of resting-state fMRI in 17 patients with AD, 17 patients with PSP, and 12 controls. Strongly connected nodes displayed more tau pathology in AD, independently of intrinsic connectivity network, validating the predictions of theories of trans-neuronal spread but not supporting a role for metabolic demands or deficient trophic support in tau accumulation. This was not a compensatory phenomenon, as the functional consequence of increasing tau burden in AD was a progressive weakening of the connectivity of these same nodes, reducing weighted degree and local efficiency and resulting in weaker ‘small-world’ properties.

Conversely, in PSP, unlike in AD, those nodes that accrued pathological tau were those that displayed graph metric properties associated with increased metabolic demand and a lack of trophic support rather than strong functional connectivity. Together, these findings go some way towards explaining why AD affects large scale connectivity networks throughout cortex while neuropathology in PSP is concentrated in a small number of subcortical structures. Further, I demonstrate that in PSP increasing tau burden in midbrain and deep nuclei was associated with strengthened cortico-cortical functional connectivity. Disrupted cortico-subcortical and cortico-brainstem interactions meant that information transfer took less direct paths, passing through a larger number of cortical nodes, reducing closeness centrality and eigenvector centrality in PSP, while increasing weighted degree, clustering, betweenness centrality and local efficiency.
Our results have wide-ranging implications, from the validation of models of tau trafficking in humans to understanding the relationship between regional tau burden and brain functional reorganization.
Chapter 3: Top-down frontal contributions to predictive processes in speech perception

3.1 Preface

Chapter 2 provided a link between the molecular pathology of neurodegeneration and the physiology of dementia in terms of functional connectivity. In this experimental chapter I explore how disruption of neural circuits gives rise to changes in core cognitive systems. To do this, I turn to a new model disease, non-fluent variant Primary Progressive Aphasia (nfvPPA). This disease is an ideal model for behavioural studies as it has a clinical phenotype that is restricted to the speech domain in early disease, allowing for precisely controlled auditory psychophysical assessment supported by well-preserved general cognition in early disease. Additionally, the focal pattern of mild atrophy in frontal lobes with relatively preserved cortical volume elsewhere greatly simplifies the analysis of neurophysiological data.

The concept for this study arose from the clinical observation that our patients with nfvPPA, who have classically been thought to have an output aphasia, were complaining of deafness unrelieved by hearing aids. This symptom had previously been validated by the observation that patients with nfvPPA perform surprisingly poorly at some tasks of basic auditory processing (Goll et al., 2010, Grube et al., 2016). However, the reasons for these deficits were poorly explained. Separately, for his PhD Ediz Sohoglu (supervised by my advisor, Dr Matt Davis) conducted a series of experiments in young healthy individuals in which he assessed the perceptual and neurophysiological impacts of manipulating prior expectations for word identity (Sohoglu et al., 2012, Sohoglu et al., 2014). He demonstrated that this manipulation modulated activity in the same frontal brain regions in which patients with nfvPPA have significant neurodegeneration, and activity in temporal brain regions, where patients do not have significant neurodegeneration. The question therefore arose as to whether abnormal, network-level neural processing of prior expectations might account for the perceptual symptoms in nfvPPA, and whether the demonstration of a diaschisis in temporal lobe activity caused by frontal neurodegeneration might provide
evidence for causal top-down frontal contributions to predictive processes in speech perception. In this chapter I primarily address the latter question, examining normative language functioning using nfvPPA as a model of abnormal network function. In chapter 5, I discuss the implications of this work for our understanding of receptive auditory difficulties in nfvPPA.

A paper based on this chapter has been published in *Nature Communications* with myself as first author (Cope *et al.*, 2017a). The study was conceptualised by my supervisor Prof James Rowe and my advisor Dr Matt Davis. Experiment one is based on a paradigm developed by Dr Ediz Sohoglu, and I am very grateful to him for allowing me to use his stimulus presentation materials. I modified these materials to ensure suitability for our participants with advice and guidance from Prof Karalyn Patterson and Mrs Kate Dawson (research nurse). I developed experiment two together with my advisor. I am grateful to Dr Helen Blank for the stimuli used in that experiment (Blank and Davis, 2016). I developed the subjective symptom rating scales with advice from Dr Bob Carlyon. I am grateful to Dr Manon Grube for allowing me to exactly replicate a subset of the experiments from her *Brain* paper (Grube *et al.*, 2016), using her computer code. I collected all of the behavioural and MEG data. I am grateful to Mrs Julie Wiggins (research nurse) for accompanying the patients to their structural MRI scans. The Bayesian perceptual modelling is based on work previously published by Dr Sohoglu (Sohoglu and Davis, 2016), developed by myself to allow for the assessment of perceptual clarity and derivation of quantitative parameters ([https://github.com/thomascope/Bayesian_Model_Code](https://github.com/thomascope/Bayesian_Model_Code)). I undertook all MEG analyses, but I was afforded a head-start by the kind sharing of computer code and detailed advice from Dr Ediz Sohoglu, Prof Richard Henson and Dr William Sedley. I updated this code for SPM12, and integrated it into a standardised pre-processing pipeline that is publically available and has now been used by others ([https://github.com/thomascope/VESPA/tree/master/SPM12version/Standalone%20preprocessing%20pipeline/seconddistrib](https://github.com/thomascope/VESPA/tree/master/SPM12version/Standalone%20preprocessing%20pipeline/seconddistrib)). I analysed the structural MRI data with guidance from Mr P Simon Jones. I am grateful to Dr Bob Carlyon, Prof Richard Henson, Prof Tim Griffiths and Dr Dennis Norris for productive external perspectives about data interpretation, and to Dr Kieren Allinson for neuropathological opinion. Finally, I am indebted to six anonymous reviewers (two at *Science* and four at *Nature*).
Communications) for helpful suggestions of refinements to the Bayesian modelling and additional MEG analyses of single subject timecourses, coherence and connectivity. I constructed all of the figures, and wrote the text with comments from all of those listed above.
3.2 Introduction

It has long been recognised that perception relies on the integration of sensory input with expectations based on prior knowledge or experience (von Helmholtz, 1925 translation of 1868 original). This can be instantiated in hierarchical generative models, which contain both top-down connections for priors or beliefs about sensory evidence, and bottom-up connections for recognition or prediction error. The layers of these hierarchical models represent progressively more abstract descriptions of the underlying sensory data (Hinton, 2007, Friston, 2008). An influential implementation of hierarchical generative models is known as predictive coding (Rao and Ballard, 1999, Friston, 2005), in which the top-down generative connections express predictions for expected sensory signals while bottom-up processes pass forward prediction errors to update the model (Figure 3-1 A). This method of information transfer is highly efficient (Musmann, 1979). Neural models of predictive coding are well formalised, and I set our results into this framework throughout the manuscript, but it should be emphasised that the results I present are equally applicable to all hierarchical generative frameworks involving top-down and bottom-up influences at each processing stage.

There is empirical evidence for predictive coding in health, for vision (Murray et al., 2002, Alink et al., 2010), hearing (Arnal et al., 2009, Grahn and Rowe, 2009, Phillips et al., 2016) and the link between perception and action in motor control (Kilner et al., 2007, Shipp et al., 2013, Wolpe et al., 2016). Furthermore, dysfunctional predictive coding mechanisms can explain a range of neurological and psychiatric phenomena, in schizophrenia (Brown et al., 2013), functional movement disorders (Edwards et al., 2012, Pareés et al., 2014), alien limb syndrome (Wolpe et al., 2014), tinnitus (Sedley et al., 2016) and hallucinations (Kumar et al., 2014). Although these disorders have been explained in terms of aberrant predictive coding, the functional consequences of degradation of the neural architecture responsible for generating top-down predictions are unknown. This is a critical and novel test for hierarchical models of perception, which motivates the following hypothesis: degeneration of top-down prediction mechanisms in frontal lobe should have a substantial impact on lower-level sensory responses in temporal lobe, and should impair perceptual function when prior knowledge and sensory input must be combined.
Figure 3-1 An illustration of the experimental motivation. A: A schematic Bayesian framework for predictive coding in speech perception. B: The putative brain basis of this framework (Park et al., 2015). Predictions are generated in inferior frontal gyrus and/or frontal motor speech regions (pink), and instantiated in auditory regions of superior temporal lobe (pale blue). C: The two dimensional experimental manipulation employed here to detect a dissociation between normal temporal lobe responses to sensory detail (number of vocoder channels) and abnormal frontal lobe responses to prior congruency. D: Our experiment relies on detecting the consequences of degraded predictions in abnormal frontal brain regions by measuring their effects in normal temporal regions. E: Voxel-based morphometry in our patient group. Regions coloured in red displayed consistent reductions in grey matter volume (FWE p<0.05). Regions coloured blue had strong evidence for normal cortical volume in nfvPPA (Bayesian probability of the null >0.7, cluster volume >1cm3). Uncoloured (grey) areas had no strong evidence for or against atrophy.

I test this hypothesis in the context of neurodegeneration of the language network from nfvPPA. Speech is a natural domain in which to study predictive coding, as humans are
able to exploit a wide variety of visual, contextual and semantic cues to improve perception, especially in difficult listening environments (Sumby and Pollack, 1954, Miller and Isard, 1963). Indeed, contradictory beliefs established by mismatching visual and auditory speech can lead to false perception (McGurk and MacDonald, 1976, Van Wassenhove et al., 2005). It is important to note that such multi-modal integration can be modelled in terms of predictive coding regardless of whether or not visual information occurs before auditory information (Schwartz and Savariaux, 2014); what is important is that auditory sensory predictions are set up based on information from prior experience, sentential context or sensory information from another domain. I exploited the importance of written text in supporting perception of degraded speech (Remez et al., 1981, Sohoglu et al., 2014) (Figure 3-1 C). There is evidence for left lateralised top-down information transfer (Park et al., 2015) from frontal language (Abdelnour et al., 2014) and motor speech (Raj et al., 2012, Raj et al., 2015) regions to auditory cortex during speech perception (Figure 3-1 B); in predictive coding theory this top-down transfer generates prior expectations for speech content and explains how listeners combine prior knowledge and sensory signals during perception and perceptual learning (Blank and Davis, 2016, Sohoglu and Davis, 2016).

To assess the effects of disrupted predictions (Figure 3-1 D) I studied patients with early non-fluent primary progressive aphasia (nfvPPA), which is associated with selective neurodegeneration of the frontal lobe language and motor speech areas (Murray et al., 2005), but preservation of temporal lobe auditory regions (Figure 3-1 E). Disordered speech output in nfvPPA is characterised by apraxia of speech and/or agrammatism (Gorno-Tempini et al., 2011). In contrast to stroke aphasia the neural damage in frontal regions is partial (Mackenzie et al., 2006, Sampathu et al., 2006, Mohandas and Rajmohan, 2009, Mackenzie et al., 2011) enabling us to study a disruption of predictive mechanisms, rather than a system reorganised following their complete absence. Additionally, this patient cohort presents fewer problems for the modelling and interpretation of magnetoencephalography (MEG) or electroencephalography (EEG), as atrophy is subtle in early nfvPPA (Gorno-Tempini et al., 2004, Rogalski et al., 2011).

For patients and matched control participants I recorded behavioural and neural data showing the influence of top-down and bottom-up manipulations on speech perception
using a well-understood paradigm involving presentation of written text that matches or mismatches with degraded spoken words (Sohoglu et al., 2012, Sohoglu et al., 2014, Blank and Davis, 2016) (Figure 3-1 C). With this paradigm, I can determine whether and how frontal cortical neurodegeneration impairs speech perception. The presence and function of top-down influences on speech perception is controversial (see (Norris et al., 2000, McClelland et al., 2006, McQueen et al., 2006, Norris et al., 2016)), as is the question of whether frontal cortical regions make a critical contribution to speech perception, through predictive coding or alternative mechanisms. Some authors suggest that these contributions are task-specific and not a core component of speech perception systems (Lotto et al., 2009). The present study provides causal neural evidence with which to assess both of these claims.

I hypothesised that, if the frontal lobes play a central role in the reconciliation of predictions with sensory evidence, frontal neurodegeneration would result in a disrupted neural effect of prior context in temporal lobe speech regions. Further, if predictive coding is a core mechanism for speech perception, then I hypothesised that frontal neurodegeneration would disrupt the application of prior knowledge, leading to aberrant speech perception in nfvPPA patients.

3.3 Methods:

All study procedures were approved by the UK National Research Ethics Service. Protocols for magnetoencephalography and magnetic resonance imaging were reviewed by the Suffolk Research Ethics Committee, and for neuropsychological tests outside of the scanner environments by the County Durham & Tees Valley Ethics Committee. All participants had mental capacity and gave informed consent to participation in the study.

3.3.1 Participants

Eleven patients with early nfvPPA were identified according to consensus diagnostic criteria (Gorno-Tempini et al., 2011). Particular care was taken to exclude patients with yes/no confusion that would confound behavioural analysis, and to include only those who lacked the lexical difficulties of logopenic and mixed aphasias, in order to
select patients most likely to have underlying tau or TDP-43 related pathology preferentially involving frontal lobes (rather than Alzheimer-type pathology of parietal lobes) (Rogalski et al., 2011, Rohrer et al., 2012, Sajjadi et al., 2012, Sajjadi et al., 2014, Mandelli et al., 2016). On the short form of the Boston Diagnostic Aphasia Examination all patients scored 10/10 for responsive naming, 12/12 for special categories, at least 15/16 for basic word discrimination and at least 9/10 for following complex commands. One patient was unable to tolerate the MEG scanner environment so, for that case, results contribute only to the behavioural analysis.

I obtained standardised volumetric T1 MRI scans on nine of the patients within two months of their MEG session, for co-registration and voxel-based morphometry.

Eleven age-, gender- and education-matched controls were recruited for behavioural and physiological studies (demographic information in Table 3-1). Thirty-six healthy age-matched control MRI datasets were selected for voxel-based morphometry.

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Age</th>
<th>Gender</th>
<th>Age leaving Education</th>
<th>MMSE</th>
<th>ACE-R</th>
<th>Raven’s Matrices</th>
</tr>
</thead>
<tbody>
<tr>
<td>nfvPPA</td>
<td>11</td>
<td>72 (9)</td>
<td>8F, 3M</td>
<td>18 (3)</td>
<td>28 (2)</td>
<td>84 (12)</td>
<td>36 (9)</td>
</tr>
<tr>
<td>MEG Controls</td>
<td>11</td>
<td>72 (8)</td>
<td>7F, 4M</td>
<td>17 (2)</td>
<td>29 (1)</td>
<td>95 (2)</td>
<td>46 (7)</td>
</tr>
<tr>
<td>MRI Controls</td>
<td>36</td>
<td>73 (7)</td>
<td>17F, 19M</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3-1 Demographic details of the experimental groups.
Mean (standard deviation). There were no statistically significant differences in age, gender or education between nfvPPA patients and MEG controls. nfvPPA patients scored more poorly than controls on the Addenbrooke’s Cognitive Examination (Revised) and Raven’s Progressive Matrices, but were still within the population normal range. Most of the difference between nfvPPA and controls on the ACE-R was accounted for by verbal fluency. Audiometric thresholds are available in Figure 3-2 B. One patient was unable to tolerate the MEG scanner environment so, for that case, results contribute only to the behavioural analysis.

3.3.2 Voxel Based Morphometry

Nine patients with nfvPPA underwent structural MR imaging at the Wolfson Brain Imaging Centre, University of Cambridge, UK using a 3T Siemens Magnetom Tim Trio scanner with a Siemens 32-channel phased-array head coil (Siemens Healthcare,
Erlangen, Germany). A T1-weighted magnetization-prepared rapid gradient-echo (MPRAGE) image was acquired with repetition time (TR)=2300ms, echo time (TE)=2.86ms, matrix=192x192, in-plane resolution of 1.25x1.25mm, 144 slices of 1.25mm thickness, inversion time=900ms and flip angle=9°. These images were compared to 36 healthy control scans with identical parameters.

All analysis was performed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm). Images were first approximately aligned by coregistration to an average image in MNI space, before segmentation and calculation of total intracranial volume (TIV). After segmentation, a study-specific DARTEL template was created from the patient scans and the 9 closest age-matched controls using default parameters. The remaining controls were then warped to this template. The templates were affine aligned to the SPM standard space using ‘Normalise to MNI space’ and the transformation applied to all individual grey-matter segments together with an 8mm FWHM Gaussian smoothing kernel. The resulting images were entered into a full factorial general linear model with a single factor having two levels, and age and TIV as covariates of no interest. This model was estimated in two ways. Firstly, a classical estimation based on the restricted maximum likelihood was performed to assess for group difference. Voxels were defined as atrophic if they were statistically significant at the peak FWE p<0.05 level. Secondly, a Bayesian estimation was performed on the same model, and a Bayesian contrast between patients and controls specified. The resulting Bayesian map was subjected to hypothesis testing for the null in SPM12, resulting in a map of the posterior probability of the null at each voxel. For visualisation in Figure 3-1 E, this map was thresholded for posterior probabilities for the null above 0.7 and cluster volumes of greater than 1cm³.

To assess for the dissociation between atrophic and preserved cortical regions, both model estimations were assessed at voxels of interest. Atrophic regions were assessed with classical statistical parametric mapping across the whole brain. This identified highly significant peaks centred in left (MNI -37, 17, 7) and right (MNI 37, 20, 6) inferior frontal regions but no peaks in superior temporal regions (supplementary table 1). Superior temporal voxels of interest were therefore defined from the Neuromorphometrics atlas, at locations corresponding to left primary auditory cortex (MNI -59, -24, 9) and superior temporal gyrus (MNI -67, -17, 3). Frequentist
probability of atrophy and Bayesian probability of no atrophy are reported at each of these four locations in results.

To create Figure 3-1 E, a rendering of the significant regions in each analysis, the DARTEL template images were further warped using the ‘Population to ICBM Registration’ function with the transformation parameters applied to all thresholded statistical maps.

To extract grey matter volume for correlation with the latency of MEG responses (Figure 3-11 E, F), a full factorial general linear model was constructed with the 9 patients alone, with age and TIV as covariates of no interest. Each subject’s age and TIV adjusted grey matter density was extracted at the voxel closest to the MEG regions of interest (left frontal [-46 2 28]; left temporal [-56 -34 12]). A secondary SPM analysis with neural latency entered as a covariate into the model and small volume correction of 8mm (to match the FWHM Gaussian smoothing kernel) at each location confirmed the results of the primary analysis – correlations below p<0.05 were observed at the frontal but not the temporal location.

3.3.3 Modifications to Sohoglu MEG paradigm

Stimuli and experimental procedures during neuroimaging were closely modelled on a task previously performed to evaluate influences of prior knowledge and sensory degradation in young, healthy listeners (Sohoglu et al., 2012) (Figure 3-3 A). In this task, individuals are presented with a written word, followed 1050 (+/-50) ms later by a spoken word, which is acoustically degraded using a noise vocoder (Shannon et al., 1995). After a further delay of 1050 (+/-50) ms, participants are asked to rate the perceptual clarity of the vocoded word. This allows for a factorial manipulation of two stimulus dimensions that in previous studies have been shown to affect speech perception: (1) the degree of correspondence between the written and spoken words can be modified by presenting text that either matches or mismatches with the speech, (2) the amount of sensory detail can be manipulated by varying the number of channels in the noise vocoder. 108 trials of each condition were presented across six blocks. Each block contained 18 trials of each combination of vocoder channel number and cue congruency in one of two fixed random orders counterbalanced across groups. To
avoid predictability, each subject observed 216 words twice in written form and twice in spoken form (once as part of a match pair and once as part of a mismatch pair) and 108 words only once (in either a match or mismatch pair). The following modifications were made to the Sohoglu et al. paradigm in order to simplify procedures for patients and elderly controls. The number of channels used in the vocoder was doubled to 4/8/16, the range expected to cover the steepest portion of the psychometric response function in older adults (Sheldon et al., 2008). The duration of the visual prime was increased from 200ms to 500ms. The resolution of the clarity rating scale was reduced from 1-8 to 1-4, so that a four-button box could be used to indicate responses. Finally, the neutral priming condition was removed to reduce the overall number of trials inside the scanner by a third, while minimising the reduction in the power with which I could test for an effect of prime congruency. 108 trials of each condition were presented across six presentation blocks. A fully crossed 2x3 factorial design was employed, with two levels of prime congruency (matching/mismatching) and three levels of sensory detail (4/8/16 vocoder channels). Each spoken word was presented no more than twice to each participant, once with a congruent prime and once with an incongruent prime.

3.3.4 Behavioural data stimuli and procedure

To ensure that I was observing an effect of prediction and that patients were not simply being confused by mismatching written words, experiment 1 was repeated outside the scanner with identical parameters but an additional, neutral, cue condition (Figure 3-8), mirroring that of Sohoglu et al. (2012). 18 trials of each condition were presented in a single block.

A second experiment was undertaken to assess participants’ ability to identify vocoded words. In this task, (Figure 3-3 C), no prior written text was provided. Participants simply heard a noise vocoded word and, 1050 +/- 50 ms after word onset, were presented with four alternatives, from which they selected the word that they had heard. I used this forced choice response format to ensure that performance was not confounded by speech production difficulties. As in experiment 1, the clarity of the spoken word was varied between 4, 8 and 16 noise vocoder channels. The closeness of the three distractor items to the correct response was also varied by manipulating the number of neighbours between the spoken word and the alternatives. In the example
shown in Figure 3-3 C, the spoken word is ‘Gaze’. The alternatives presented to the participant comprised ‘Daze’ (an offset neighbour), ‘Gaze’ (the target), ‘Game’ (an onset neighbour), and ‘Then’ (not a neighbour). This set of four response alternatives occurred in trials where the spoken word was either: (1) “Gaze” (such that there were two word neighbours, “Daze” and “Game” in the response array), (2) “Daze” (only one onset neighbour, “Gaze”), (3) “Game” (one offset neighbour “Gaze”), or (4) “Then” (no neighbours in the response array). This achieved a factorial experimental design with a 3-level manipulation of sensory detail fully crossed with a 4-level manipulation of distractor difficulty. There were 90 trials in a single block; at each of the 3 levels of sensory detail there were 10 spoken words in each of cases (1) and (4) above, and 5 each in (2) and (3). This resulted in 30 sets of four response options each being presented three times and having 3 of its 4 members heard during the experiment. Word presentation orders were randomised across participants, but the sensory detail and written neighbour difficulty order was fixed.

Auditory stimuli were presented in a quiet room through Sennheiser HD250 linear 2 headphones, driven by a Behringer UCA 202 external sound card, and visual stimuli were displayed on a laptop computer screen. Participants indicated responses either by pressing a number on a keyboard (clarity rating outside of MEG) or a button on a custom made response box (all other experiments).

3.3.5 Behavioural data modelling

Subjective ratings of clarity were modelled using an hierarchical Bayesian inference approach previously described for data of this type (Sohoglu and Davis, 2016) (Figure 3-9). This model exploits the principles of predictive coding, in which perception arises from a combined representation of sensory input and prior beliefs (Rao and Ballard, 1999, Friston, 2005, Hickok, 2012). It is able to explain both the perceptual benefit of matching prior information, as the precision of the ‘posterior’ (or subjective experience) is increased by congruency between the prior and the sensory input, as well as the previously observed dissociated modulations in superior temporal gyrus activity by cue congruency and sensory detail (Sohoglu et al., 2012, Sohoglu and Davis, 2016).
The model is able to predict subjective clarity as a function of the precision of the posterior distribution, which is estimated as the precision of the sensory input multiplied by a weighted function of the precision of the prior. The weighting given to the prior information depends on its congruency. In the case of a mismatching prior, the precision of the posterior simply matches the precision of the sensory input, while for matching priors it increases as a function of the precision of prior expectation. Finally, the precision of the posterior is compared against a perceptual threshold (below which degraded speech is deemed completely unclear and given a rating of 1), and the height above this threshold is mapped to the rest of the rating scale (subjects were instructed to use the full range of the rating scale, and undertook practice trials to familiarise themselves with the range of experimental stimuli – all subjects were able to do this).

The precision of the sensory input was individually pre-defined for each subject at each level of sensory detail, based on their measured ability to correctly report words with that number of vocoder channels (Figure 3-3 D). The weighting of matching prior information was defined as 0.5, reflecting the experimental context in which 50% of written words were congruent with (i.e. matched) the degraded spoken words. In ‘open set’ listening situations like those used in the present experiment, the weighting of prior expectation for the occurrence of any uncued word is the inverse of the number of nouns in the participant’s lexicon. I therefore approximated the weighting to zero for both mismatching and neutral prior expectations (Blank and Davis, 2016). This accounts for the observation that clarity ratings in this and previous experiments were almost identical following mismatching and neutral (uninformative) written words (Sohoglu et al., 2012, Sohoglu et al., 2014).

It is possible that the patients might have an inappropriately high ‘weighting’ for the written cue. In other words, they apply their prior expectations ‘inflexibly’, being unable to account for the fact that they will only be correct half the time. As we have a binary situation (the text was either fully matching or fully mismatching) I am unable to assess this possibility directly – in my model allowing the weighting of matching prior information to vary would be mathematically equivalent to allowing the precision of the prior to vary. Accordingly, I use the terms ‘excessively precise priors’ and ‘inflexible priors’ interchangeably.
The precision of the prior expectation and the level of the perceptual threshold were then individually optimised to provide the best-fit to each subject’s clarity ratings by global minimisation of squared residuals (using the Global Optimisation Toolbox in MATLAB).

### 3.3.6 Alternative data models

It was observed that model fits were less good for the mismatch condition in some controls, with the model systematically under-predicting slope. The primary driver of this effect seemed to be a ‘washing out’ of the effect of prior knowledge in the face of very clear speech for some individuals (i.e. there is less perceptual clarity benefit to prior knowledge if the auditory token is itself very clear). This has been previously observed in young healthy individuals (Sohoglu et al., 2012, Sohoglu et al., 2014) and seems to result from participants nearing the upper end of their psychometric response functions and entering a region of non-linear response. That is, doubling from 8 to vocoder 16 channels results in more benefit in terms of perceptual clarity than doubling from 16 to 32 channels. Indeed, as can be observed from Figure 3-3 D many controls are reaching ceiling performance at reporting 16 channel vocoded speech in the absence of prior information. While we know that humans are able to detect sensory detail changes even in the face of unimpaired word identification (think, for example, of hi-fi reviews), it is not unreasonable to assume that these changes in perceptual clarity might result in smaller rating changes than those that meaningfully improve word identification performance.

The model as originally formulated does not account for non-linearities of this type because the relative increase in the precision of the posterior distribution resulting from congruent prior expectations is modelled as being a constant function of sensory detail. In reality, congruent prior expectations are shifting the position on the psychometric response function (see (Sohoglu et al., 2014)). As I do not have experimental data for higher degrees of sensory detail (for example unprimed 24 or 32 channel speech) I cannot account for this directly. If, however, I allow the weighting of prior expectations to vary in the Match case for 16 channel speech this would simulate the effect of entering a flatter portion of the psychometric response function. Doing this did improve the model fits overall by reducing the under-prediction of slope in the
controls. Although this resulted in small changes in the optimised values of the other model parameters it did not affect any of the statistical results or relationships reported in the chapter (the priors are still more precise in the patients at p<0.01, and their precision still significantly correlates with beta power during the instantiation of predictions with r=-0.51).

I compared the performance of the original model and the new model with the Akaike information criterion (corrected for small samples, AICc), which assesses whether additional model parameters sufficiently improve the information provided to account for reductions in parsimony. The simpler (original) model was favoured for 10 of the 11 patients, and 8 of the 11 controls. On average, the simpler model had an AICc 4.02 points lower than the more complex model (lower AICc scores are better).

Therefore, I retain the original model in the results section. Overall, however, I take reassurance from the fact that both models result in the same statistical relationships between the precision of prior expectations and beta power during the instantiation of predictions.

3.3.7 MEG and EEG data acquisition and analysis

An Elekta Neuromag Vectorview System was used to simultaneously acquire magnetic fields from 102 magnetometers and 204 paired planar gradiometers, and electrical potentials from 70 Ag-AgCl scalp electrodes in an Easycap extended 10-10% system, with additional electrodes providing a nasal reference, a forehead ground, and paired horizontal and vertical electrooculography. All data were digitally sampled at 1kHz and high-pass filtered above 0.01Hz. Head shape, EEG electrode locations, and the position of three anatomical fiducial points (nasion, left and right pre-auricular) were measured before scanning with a 3D digitiser (Fastrak Polhemus). The initial impedance of all EEG electrodes was optimised to below 5 kiloOhms, and if this could not be achieved in a particular channel, or if it appeared noisy to visual inspection, it was excluded from further analysis.

During data acquisition, the 3D position of five evenly distributed head position indicator coils was monitored relative to the MEG sensors (magnetometers and gradiometers). These data were used by Neuromag Maxfilter 2.2, to perform Signal
Source Separation (Taulu et al., 2005) for motion compensation, and environmental noise suppression.

Subsequent pre-processing and analysis was undertaken in SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK), FieldTrip (Donders Institute for Brain, Cognition, and Behavior, Radboud University, Nijmegen, The Netherlands) and EEG lab (Swartz Center for Computational Neuroscience, University of California San Diego), implemented in MATLAB 2013a. Artefact rejection for eye movements and blinks was undertaken by separate independent component analysis decomposition for the three sensor types. For MEG data, components were automatically identified that were both significantly temporally correlated with contemporaneous electrooculography data and spatially correlated with separately acquired template data for blinks and eye movements. For EEG data, components spatially and temporally consistent with eye blinks were automatically identified with ADJUST (Mognon et al., 2011). These components were then projected out from the dataset with a translation matrix. Due to a technical difficulty during acquisition, one control subject had no signal recorded from two thirds of their EEG sensors, and one patient had seven sensors that failed quality control – these individuals were excluded from the EEG analysis, but included in MEG and behavioural analyses.

For evoked analysis, the data were then sequentially epoched from -500 to 1500ms relative to speech onset, downsampled to 250Hz, EEG data referenced to the average of all sensors, baseline corrected to the 100ms before speech onset, lowpass filtered below 40Hz, robustly averaged across epochs, refiltered below 40Hz (to remove any high frequency components introduced by the robust averaging procedure), planar gradiometer data were root-mean-square combined, all data were smoothed with a 10mm spatial kernel and 25ms temporal kernel before conversion to images in a window from -100 to 900ms for statistical analysis.

For induced analysis, the de-artefacted continuous data were downsampled, re-referenced, baseline corrected as above, lowpass filtered below 100Hz, notch filtered to exclude line noise between 48 and 52Hz, then epoched, before being submitted to four separate time frequency decompositions by the Morlet wavelet method: two separate time windows of -500 to 1500ms relative to written word and speech onsets were examined, with and without pre-subtraction of the condition-averaged waveform from
every trial. These were robustly averaged and log rescaled compared to pre-visual baseline power in each frequency band. Morlet decomposition parameters were focused for sensitivity to low-mid frequencies, with 7 wavelet cycles in a range from 4 to 80Hz in steps of 2Hz.

3.3.8 Sensor-space evoked analysis

For each sensor type separately, a flexible factorial design was specified in SPM12, and interrogated across all participants for main effects of prime congruency and clarity. For all sensor types, a scalp position of peak effect was defined where peak FWE \( p < 0.01 \). The sensor data at this scalp position was then compared across groups at every time point. A significant group x condition interaction was defined as at least 7 consecutive timepoints of \( p < 0.05 \), exceeding the temporal smoothing induced by lowpass filtering at 40Hz. This approach does not represent double-dipping as the location of interest was defined by an orthogonal contrast (Friston and Henson, 2006, Kriegeskorte et al., 2009, Kilner, 2013), and in any case for the effect of congruency (where group x condition interactions were observed with this method), for both the planar gradiometers and the magnetometers the location of peak effect for patients alone was within 2mm of the conjoint peak effect.

3.3.9 Evoked data source reconstruction

Source inversion methods by the sLORETA algorithm were identical to those employed by Sohoglu et al. (2012), except that they were undertaken in SPM12 rather than SPM8. It was observed that the time windows of interest defined in healthy young controls were slightly earlier than the main data features in my cohort of more elderly controls, who displayed similar overall profiles to patients with nfvPPA (Figure 3-4). The time windows of interest were therefore slightly lengthened and delayed, to ensure that the main data variance was captured.

The aim of the source data analysis was to localise and explore the brain basis of the group by congruency interaction statistically demonstrated in the sensor space data. While localisation of the main effect of clarity was undertaken across all individuals,
and is shown in Figure 3-7, in the absence of a group by clarity interaction in sensor space, no further analysis was performed on this condition.

From the previous studies in healthy young controls, it was anticipated that significant main effects of congruency would be observed in opposite directions in left frontal regions and left superior temporal gyrus. This was indeed the case, with a small left frontal region being significantly more active across the whole time window with Matching prior information, and a larger region centred on left superior temporal gyrus being significantly more active with Mismatching prior information. These peak locations were defined as voxels of interest, and the source power averaged for each condition at each location within every time window of interest for every individual. Independent, repeated measures ANOVAs were then performed in each time window. Those that demonstrated a statistically significant main effect of group or a group by congruency interaction are illustrated in Figure 3-7 C (the main effect of congruency was not examined, as this would represent double dipping at these voxels).

3.3.10 Sensor-space induced analysis

The primary analysis of induced data was undertaken in the planar gradiometers because of their superior signal to noise ratio for data of this kind (Muthukumaraswamy and Singh, 2013). Other sensor types were examined secondarily to check for consistency of effect, which was confirmed in all cases. Visual inspection of the time x frequency data at a variety of scalp positions revealed no clear difference in the pattern of effect (although its strength differed, as shown in Figure 3-12), so data were collapsed across all sensors for statistical comparison. A flexible factorial design was specified in SPM12 for time x frequency data across a time window of -100 to 1000ms, and interrogated across all participants for all contrasts of interest (main effects of group, prime congruency and sensory detail, all pairwise and the three-way interaction). A second, confirmatory, analysis was performed with the condition-averaged waveform subtracted from every trial with identical statistical results, demonstrating that the effects were induced rather than evoked (I make no claims as to whether they are dynamic or structural (David et al., 2006)).
3.3.11 Induced data source reconstruction

The significant group by condition interactions observed in alpha and beta frequency bands were localised with the ‘Data Analysis in Sensor Space’ toolbox in SPM12. sLORETA was not available in this toolbox, so for closest comparability with the evoked reconstructions, the eLORETA algorithm was used. Reconstructions used time frequency data at the frequency of maximum group x congruency interaction, +/-6Hz. Data were truncated at 50 principal components, to avoid any problems with beamforming after Signal Source Separation, which reduces the number of independent components in the data to around 70 (Taulu et al., 2005). Lead fields were calculated over a window of interest from 350ms to 900ms, and sources reconstructed in three separate time windows of equal duration defined by the sensor space group x congruency interaction: 300-450ms, where controls had a greater main effect of congruency; 450-600ms, where there was no group x congruency interaction; and 600-750ms, where patients had a greater main effect of congruency. A flexible factorial design was specified in SPM12, and the group by congruency interaction (already statistically demonstrated across the whole brain) thresholded for visualisation at uncorrected p<0.01.

3.3.12 Coherence and Connectivity Analyses

The timeseries of the frontal ([46,2,28]) and temporal ([56,-34,12]) sources of interest (see Figure 3-7 B) were extracted between 0 and 912ms after every spoken word using the function spm_eeg_inv_extract. The condition-averaged waveform (i.e. the evoked response) in each source was then subtracted from every trial to result in data with zero-mean and approximate stationarity within the time window of interest. The Fourier spectra were then computed in FieldTrip using multitapers with a +/- 4Hz smoothing box. This decomposition was then subjected to separate FieldTrip connectivity analyses with either imaginary coherence or Granger causality (based on an optimised multivariate auto-regressive model of maximum order 10 (Harrison et al., 2003, Schnitzler and Gross, 2005)). This same procedure was repeated 1000 times with the trial labels in each region shuffled to create a null distribution. Statistical assessment of the presence of coherence or connectivity at each frequency involved the
comparison of the observed data against the null distribution (Figure 3-13 A, C). Between-group comparisons of imaginary coherence employed unpaired t-tests with unequal variance (the normality assumption was not violated), cluster corrected for multiple comparisons (Figure 3-13 B). To compare the strength of Granger causal relationships between regions, I first corrected for differences in signal to noise ratio between subjects and regions by dividing the magnitude of each frequency value by the across-frequency mean for that individual-region pair. This created a profile of relative influence for each region at each frequency, corrected for overall differences in signal strength. At each frequency, the significance of ‘directionality’ (i.e. temporal to frontal vs frontal to temporal) was assessed with a repeated measures general linear model, and the output corrected for multiple comparisons (Figure 3-13 D).

3.4 Results:

3.4.1 Structural consequences of nfvPPA

To confirm the dissociation between frontal atrophy and intact temporal cortex, upon which my experiment relies, I used voxel-based morphometry to compare grey matter volume in my patients with nfvPPA to 36 age-matched healthy individuals. As anticipated, brain regions of interest displayed a localised pattern of atrophy in nfPPA (Figure 3-1 E), with grey matter volume loss in left inferior frontal regions (whole brain peak FWE p = 0.001 at MNI [-37, 17, 7]; Bayes posterior probability of no difference <0.00001), but not in left primary auditory cortex (FWE p = 1; Bayes posterior probability of no difference 0.75 at MNI [-59, -24, 9]) or superior temporal gyrus (FWE p = 1; Bayes posterior probability of no difference 0.91 at MNI [-67, -17, 3]). Significant atrophy was also observed in right inferior frontal regions (FWE p = 0.004; peak MNI [37, 20, 6]) but not right primary auditory cortex (FWE p=1 at MNI [59, -24, 9]) or superior temporal gyrus (FWE p = 1 at MNI [67 -17 3]). Significant atrophy in left inferior frontal regions lay within pars triangularis, pars opercularis and anterior insula in the Desikan-Killiany Atlas (see Table 3-2 for significantly atrophic clusters and their probabilistic centre of mass based on the Neuromorphometrics atlas).
Table 3-2 Voxel-based morphometry regions of significant atrophy in nfvPPA. Corresponding to Figure 3-1 E. Output of an SPM thresholded at peak FWE p<0.05. To avoid an over-long table, only clusters of >= 500 voxels are listed here.

<table>
<thead>
<tr>
<th>cluster voxels</th>
<th>peak p (FWE)</th>
<th>peak t-score</th>
<th>MNI Co-ordinates</th>
<th>Neuromorphometrics region</th>
</tr>
</thead>
<tbody>
<tr>
<td>751</td>
<td>0.001</td>
<td>6.95</td>
<td>-14   25 53</td>
<td>Left Superior Frontal Gyrus</td>
</tr>
<tr>
<td>1675</td>
<td>0</td>
<td>6.92</td>
<td>11    8   58</td>
<td>Right Supplementary Motor Cortex</td>
</tr>
<tr>
<td>0.001</td>
<td>6.52</td>
<td>5</td>
<td>-1    49</td>
<td>Right Supplementary Motor Cortex</td>
</tr>
<tr>
<td>0.001</td>
<td>6.5</td>
<td>4</td>
<td>0     62</td>
<td>Right Supplementary Motor Cortex</td>
</tr>
<tr>
<td>0.001</td>
<td>6.65</td>
<td>-35</td>
<td>17    7</td>
<td>Left Frontal Operculum</td>
</tr>
<tr>
<td>0.021</td>
<td>5.46</td>
<td>-49</td>
<td>19    12</td>
<td>Left Opercular Part of Inferior Frontal Gyrus</td>
</tr>
<tr>
<td>0.002</td>
<td>6.49</td>
<td>31</td>
<td>-4    56</td>
<td>Right Precentral Gyrus</td>
</tr>
<tr>
<td>0.018</td>
<td>5.53</td>
<td>28</td>
<td>2     49</td>
<td>Right Middle Frontal Gyrus</td>
</tr>
<tr>
<td>0.018</td>
<td>5.52</td>
<td>19</td>
<td>0     58</td>
<td>Right Superior Frontal Gyrus</td>
</tr>
<tr>
<td>0</td>
<td>1069</td>
<td>6.27</td>
<td>40    5   9</td>
<td>Right Central Operculum</td>
</tr>
<tr>
<td>0.004</td>
<td>6.05</td>
<td>37</td>
<td>20    6</td>
<td>Right Frontal Operculum</td>
</tr>
</tbody>
</table>

3.4.2 Subjective speech perception symptoms in nfvPPA

While the core symptoms in nfvPPA relate to apraxia of speech and agrammatism, patients often complain of a feeling of ‘speech deafness’. To test for this symptom in my cohort, I asked patients and controls to rate their subjective difficulty with five listening scenarios, by placing a mark on a line from ‘very easy’ to ‘very difficult’. Patients could respond appropriately to such rating scales. Patients and controls displayed very similar mean and standard deviation of the subjective difficulty ratings for ‘speech in noise’, ‘localising sounds’, ‘understanding station announcements’ and ‘how loud others say their television is’ (all p>0.3, Figure 3-2 A). However, there was a difference in their assessments of difficulty in understanding speech in quiet environments (t(20)=2.66, p=0.015): controls universally rated this as very easy, while patients rated it almost as difficult as understanding speech in noise (interaction
F(1,20)=8.21, p=0.010). Patients and controls had similar, age-appropriate, hearing as measured by pure tone audiometry (Figure 3-2 B).

Figure 3-2 Auditory symptoms in nfvPPA
A: Self-rated listening difficulty for patients and controls for four different listening scenarios, and the related question: “how loud do people tell you your TV is?” B: Pure tone audiograms for each individual participant.

3.4.3 Evoked neural responses during the reconciliation of predictions, and their cortical origins

My hypothesis that nfvPPA patients would display aberrant effects of prior knowledge during speech perception reflects an hierarchical generative model for speech perception in which frontal brain regions both instantiate and reconcile predictions
that act top-down on temporal lobe regions (Abdelnour et al., 2014, Park et al., 2015, Raj et al., 2015). Neural responses evoked by speech presentation reflect the operation of bottom-up prediction errors in response to speech signals (Figure 3-1 A). In this framework, frontal lobe neurodegeneration in nfvPPA would be expected to alter the way that neural activity in (anatomically and functionally in-tact) temporal lobe is modulated by the fulfilment or violation of sensory predictions.

To assess the neural correlates of degraded predictive mechanisms in nfvPPA (Figure 3-1 D), I recorded simultaneous MEG and EEG during a speech perception task, in which I manipulated prior expectations using matching or mismatching text cues, before subjects heard spoken words that were varied in sensory detail by manipulating the number of vocoder channels (Figure 3-3 A) (Shannon et al., 1995). Overall evoked power in the planar gradiometers was similar for the two groups of participants, with the magnetometers displaying slightly lower overall power in patients and the EEG electrodes displaying the opposite pattern (Figure 3-4). Thus, frontal neurodegeneration does not lead to any large difference in the magnitude of the neural response evoked by single spoken words that could manifest as spurious group by condition interactions in neural activity.
Figure 3-3 Behavioural results
A: Experiment 1 design. A Match trial is shown. In a mismatch trial, the written and vocoded words would share no phonology (for example the written cue ‘clay’ might be paired with the vocoded word ‘sing’). B: Group averaged clarity ratings for each condition. Error bars represent standard error across individuals within each group. C: 4 alternative forced choice vocoded word identification task. D: Group averaged percent correct report for each condition. Chance performance at 25%. Error bars represent standard error across individuals within each group. E: Overall group fits for single subject Bayesian data modelling of the data from panel B. F: Derived parameters from the Bayesian data modelling. A.U.: Arbitrary units. Patients with nfvPPA displayed significantly more precise prior expectations than controls (Wilcoxon U(11,11) p<0.01). They also displayed a trend towards a reduction in perceptual thresholds (Wilcoxon U(11,11) p=0.075).
Figure 3-4 Total evoked power.
Root mean square averaged evoked response across all sensors for each modality, after averaging across conditions and participants. The time windows over which the evoked brain sources were reconstructed are depicted by the areas shaded in grey (90-150ms, 200-280ms, 290-440ms, and 450-700ms). Scalp topographic plots display the average evoked response within each window.
To confirm whether patients have normal responses to manipulations of sensory detail independent of predictions (as expected given preserved cortical volume in auditory cortex and superior temporal gyrus, Figure 3-1 E), I first assessed the neural effect of the number of vocoder channels (Figure 3-5). Across all individuals, a peak effect was observed in the planar gradiometers at 96ms (FWE p<0.001), in magnetometers at 188ms (FWE p<0.001) and again at 380ms (FWE p<0.001), and in EEG at 392ms (FWE p<0.001). These findings are consistent with previous studies in young individuals (Sohoglu et al., 2012). Neither at the scalp locations of the peak main effect, nor with a secondary SPM analysis of all scalp-time locations, nor with time-frequency analysis, were reliable group by sensory detail interactions demonstrated at any timepoint. Together, these results are consistent with the idea that patients and controls produce similar neural responses to manipulations of sensory detail in degraded speech. Crucially for what follows, the latency of these responses was also the same in both groups. In magnetometers, where the late effects of vocoder channel number were most clearly seen, control peaks occurred at 196ms and 340ms, and patient peaks at 172ms and 372ms, with very close agreement of waveform morphology at the scalp location of peak conjoint effect (Figure 3-5).
**Main Effect of Vocoder Channels**

<table>
<thead>
<tr>
<th>Gradiometers</th>
<th>Magnetometers</th>
<th>EEG</th>
</tr>
</thead>
<tbody>
<tr>
<td>80-96ms</td>
<td>312-550ms</td>
<td>350-500ms</td>
</tr>
</tbody>
</table>

**Figure 3-5 The effect of clarity.**

Illustrative topographic plots are shown of the main effect of vocoder channels across all participants. No group by sensory detail interactions were observed either at the peak location or in a confirmatory SPM analysis.

Having performed these control analyses, I tested my primary hypothesis by examining the neural effect of manipulating whether prior expectations matched sensory input. During the reconciliation of predictions there were significant main effects of cue congruency in all sensor types (Figure 3-6 A).
Figure 3-6 The evoked effect of cue congruency.
A: Illustrative scalp topographic plots of the main effect of cue congruency for each group from 400ms to 700ms, a period of where both groups showed a large statistical effect of congruency with similar topography. White stars indicate the scalp location of the peak congruency effect across both groups between -100ms and 900ms (FWE p<0.001 for all sensor types). B: Significant group by congruency interactions (p<0.05 sustained for more than 25ms at the scalp locations marked by white stars in the upper panel) were observed in planar gradiometers and magnetometers, and are shaded in lilac. C: Topographic plots for each group are shown averaged across each significant cluster of group by congruency interaction.

Across all individuals (controls and nfVPPA), the peak effect was observed in the planar gradiometers at 464ms (FWE p<0.001), in magnetometers at 400ms (FWE p<0.001), and in EEG at 700ms (FWE p<0.001). At these scalp locations, significant group by
congruency interactions were observed in the planar gradiometers and in the magnetometers, but not in the EEG electrodes. In the planar gradiometers, between 264ms and 464ms controls had a significantly larger effect of congruency than patients (Figure 3-6 B). The scalp topography averaged across this time window resembled that observed during the peak of the main effect, but with the pattern being stronger in controls (Figure 3-6 C). In the magnetometers, a cluster with similar timing and scalp topography was observed between 240 and 560ms (Figure 3-6 B). Two additional clusters were also observed. In later time windows, from 728 to 808ms, group by congruency interactions were observed in the opposite direction, with patients showing a significantly greater effect than controls. Again, the scalp topographies in this cluster resembled those during the main effect (Figure 3-6 C). This indicates that the effect of congruency was present in both groups, but that the effect was significantly delayed in nfvPPA. Finally, an early cluster was observed between 152ms and 224ms, with the controls displaying a significantly greater effect of congruency than patients. Intriguingly, the scalp topography in this time window was different to that observed during the conjoint main effect, with dipoles having a much more anterior centre of mass (Figure 3-6 C). This anterior topography is consistent with a frontal source, expected to appear in this earlier time window as shown in similar previous studies with young healthy listeners (Sohoglu et al., 2012, Sohoglu and Davis, 2016).

To assess the underlying neural sources of these effects, multimodal sensor data were combined (Henson et al., 2009) and inverted into source space with sLORETA (Pascual-Marqui, 2002). For the main effect of vocoder channel number, reconstructions were performed across all individuals combined, because no group difference or group by clarity interaction was demonstrated in sensor space. The main effect of sensory detail shown in MEG sensors and EEG electrodes is explained by significantly more activity for 16 channel speech in temporal lobe auditory areas in mid-latency time windows (Figure 3-7 A), replicating previous findings in younger individuals (Sohoglu et al., 2012, Sohoglu and Davis, 2016).
To localise the group by cue congruency interaction, I first display source reconstructions for the main effect of cue congruency in each group (Figure 3-7 B) for time windows defined by the main data features in overall sensor power averaged over conditions and participants replicating the methods of Sohoglu et al. (2012) (Figure 3-4). Given my findings of delayed congruency effects in patients, an additional, late, time window (710-850ms) was also examined post hoc. I focus on the two principal sources observed in young healthy individuals for this task (Sohoglu et al., 2012, Sohoglu and Davis, 2016), extracting average power in each condition at frontal and
superior temporal voxels of interest defined by the main effect of congruency averaged across the whole epoch (Figure 3-7 D). No group differences were demonstrated in the earliest (90-150ms) or latest (450-700ms) time windows. Between 200 and 280ms, there were significant main effects of group in both frontal and temporal voxels, with greater responses in controls than patients. Between 290 and 440ms, this main effect had dissipated, but there was a group by condition interaction, with controls showing a greater effect of cue congruency in the superior temporal but not the frontal voxel.

The analysis across the whole epoch is of particular interest. Across all individuals, a repeated-measures ANOVA (Table 3-3) confirmed the pattern of opposing effects of prior knowledge in frontal and superior temporal regions seen in a previous study (Sohoglu et al., 2012) (F(1,134)=60.1, p<0.001). However, there was a group by source by congruency interaction (F(1,134)=11.9, p=0.001), primarily driven by the absence of a significant effect of congruency in frontal regions in nfvPPA (Figure 3-7 B). When the total power in the frontal region was examined, a main effect of group was observed such that patients had significantly more frontal power than controls, but their modulation of frontal power by congruency was absent (Figure 3-7 D).

<table>
<thead>
<tr>
<th>Test</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1,134</td>
<td>1.1</td>
<td>0.317</td>
</tr>
<tr>
<td>Source</td>
<td>1,134</td>
<td>2003.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Condition</td>
<td>1,134</td>
<td>5.1</td>
<td>0.037</td>
</tr>
<tr>
<td>Vocoder Channels</td>
<td>2,134</td>
<td>0.0</td>
<td>0.970</td>
</tr>
<tr>
<td>Individual (Group)</td>
<td>18,134</td>
<td>1.1</td>
<td>0.383</td>
</tr>
<tr>
<td>Group*Source</td>
<td>1,134</td>
<td>0.1</td>
<td>0.816</td>
</tr>
<tr>
<td>Group*Condition</td>
<td>1,134</td>
<td>0.4</td>
<td>0.550</td>
</tr>
<tr>
<td>Group*Vocoder Channels</td>
<td>2,134</td>
<td>0.8</td>
<td>0.433</td>
</tr>
<tr>
<td>Source*Condition</td>
<td>1,134</td>
<td>60.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Source*Vocoder Channels</td>
<td>2,134</td>
<td>0.1</td>
<td>0.884</td>
</tr>
<tr>
<td>Condition*Vocoder Channels</td>
<td>2,134</td>
<td>1.8</td>
<td>0.165</td>
</tr>
<tr>
<td>Group<em>Source</em>Condition</td>
<td>1,134</td>
<td>11.8</td>
<td>0.001</td>
</tr>
<tr>
<td>Source* Individual (Group)</td>
<td>18,134</td>
<td>76.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Condition * Individual (Group)</td>
<td>18,134</td>
<td>4.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Vocoder Channels * Individual (Group)</td>
<td>36,134</td>
<td>1.2</td>
<td>0.227</td>
</tr>
</tbody>
</table>

Table 3-3 Repeated measures ANOVA of evoked source power in frontal and temporal voxels of interest over the whole post auditory epoch. Bold, red rows indicate statistically significant results.
3.4.4 Behavioural Experiment 1: Vocoder Word Clarity Rating

Given these neural differences, I sought to understand the perceptual correlates of neural delay in the reconciliation of predictions by examining the behavioural consequences of manipulations of prior knowledge in my two groups. All individuals reported that the perceptual clarity of vocoded words was significantly increased by matching text cues (Figure 3-3 B), but this effect was greater in patients with nfvPPA than in controls. A repeated-measures ANOVA (Table 3-4) revealed that Group, Number of Vocoder Channels and Cue Congruency were significant either as main effects or as part of two-way or three-way interactions.

<table>
<thead>
<tr>
<th>Test</th>
<th>df</th>
<th>F</th>
<th>p (Greenhouse-Geisser)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1,40</td>
<td>7.7</td>
<td>0.011</td>
</tr>
<tr>
<td>Clarity</td>
<td>2,40</td>
<td>69.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Group * Clarity</td>
<td>2,40</td>
<td>17.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Congruency</td>
<td>1,40</td>
<td>1.2</td>
<td>0.287</td>
</tr>
<tr>
<td>Group * Congruency</td>
<td>1,40</td>
<td>13.2</td>
<td>0.002</td>
</tr>
<tr>
<td>Clarity * Congruency</td>
<td>2,40</td>
<td>2.3</td>
<td>0.140</td>
</tr>
<tr>
<td>Group * Clarity * Congruency</td>
<td>2,40</td>
<td>8.2</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Table 3-4 Repeated Measures ANOVA of behavioural data from experiment 1.

A replication experiment outside the MEG scanner confirmed that the difference between match and mismatch trials was due to a facilitatory effect of matching prior knowledge and not simply increased confusion in the face of mismatching priors: ratings of perceptual clarity after a mismatching text cue were not statistically different from those after a ‘neutral’ or uninformative cue. Furthermore, patients with nfvPPA had a much larger difference in clarity rating between ‘neutral’ and ‘match’ trials than controls (Figure 3-8).
Figure 3-8 Replication of experiment 1 with additional, neutral cues.  
*Group averaged clarity ratings from the repetition of experiment 1 (Figure 3-3 A) with the addition of a neutral written cue (a row of “XXXX”).*

It is important to note that participants were explicitly instructed to rate clarity across their own range of perceptual experience within the experiment, and were given training until they were able to do this. Comparing clarity ratings across groups is not, therefore, a direct measure of comparative listening difficulty as a rating of ‘1’ simply means ‘one of the least clear words I heard in the experiment’, while a ‘4’ means ‘one of the clearest words I heard’. To fully assess the perceptual basis of my findings, with a further experiment and Bayesian modelling I assessed the elements contributing to perceptual clarity: 1) patients’ and controls’ ability to identify degraded spoken words and 2) participants’ introspective ability to perform higher level estimation of the global precision of sensory input.

**3.4.5 Behavioural Experiment 2: Vocoded Word Identification**

To ensure that my finding was not a consequence of impaired word identification in patients leading to a group difference in reliance on prior knowledge (Shankweiler *et al.*, 2008), I performed a second experiment in which participants identified noise vocoded words in the absence of prior expectations (Figure 3-3 C). To reduce response demands for patients with non-fluent speech I used a 4-alternative forced-choice identification task. All individuals with nfvPPA were above chance at identifying even the most degraded vocoded speech however, as a group, they did not perform quite as well as controls (Figure 3-3 D, Table 3-5). Both groups were influenced in the same way by the number of noise vocoder channels and the number of close distractor items presented as alternatives in the forced choice (Table 3-5). As expected, it was easier for
all individuals to identify words with more vocoder channels and if there were fewer close distractor items. This effect was strongest for the most degraded speech, manifesting as an interaction between vocoder channels and distractor difficulty.

<table>
<thead>
<tr>
<th>Test</th>
<th>df</th>
<th>F</th>
<th>p (Greenhouse-Geisser)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>1,120</td>
<td>10.9</td>
<td>0.003</td>
</tr>
<tr>
<td>Vocoder Channels</td>
<td>2,120</td>
<td>5.79</td>
<td>0.015</td>
</tr>
<tr>
<td>Distractor Difficulty</td>
<td>3,120</td>
<td>0.839</td>
<td>0.457</td>
</tr>
<tr>
<td>Group * Clarity</td>
<td>2,120</td>
<td>0.169</td>
<td>0.765</td>
</tr>
<tr>
<td>Group * Distractor Difficulty</td>
<td>3,120</td>
<td>1.14</td>
<td>0.334</td>
</tr>
<tr>
<td>Vocoder Channels * Distractor Difficulty</td>
<td>6,120</td>
<td>4.23</td>
<td>0.004</td>
</tr>
<tr>
<td>Group * Vocoder Channels * Distractor Difficulty</td>
<td>6,120</td>
<td>1.76</td>
<td>0.149</td>
</tr>
</tbody>
</table>

Table 3-5 Repeated measures ANOVA of behavioural data from experiment 2.

Crucially, these data show that a lower-level impairment in perceiving vocoded speech cannot be the sole explanation of my finding of an increased congruency effect in nfvPPA patients. Although they did not perform as well at identifying speech of equivalent sensory detail, these patients performed better at identifying speech with 8 channels than controls did with 4 channels. Yet, patients still display a much larger congruency effect for 8-channel vocoded words than controls do for 4-channel speech. Hence, the magnitude of congruency effects in clarity rating is not simply related to objective abilities at word identification, but rather reflects a difference in the mechanisms by which prior knowledge influences lower-level perceptual processing. I investigate the nature of this effect with a Bayesian perceptual model combining word report and clarity rating data.

3.4.6 Bayesian modelling of experiments 1 and 2

To dissociate changes in the precision of predictions from difficulties with higher level estimation of the precision of sensory input, I performed hierarchical Bayesian inference simulations (c.f. (Sohoglu and Davis, 2016); Figure 3-9).

Individual differences in word discriminability were accounted for by defining the precision of sensory input for each subject as the percentage above chance for word identification at each vocoder channel number in experiment 2. This allowed us to
individually optimise two free parameters against the clarity ratings measured in experiment 1. These parameters were the precision of prior expectations (as measured by their standard deviation), and a perceptual threshold below which the observer rated speech as unclear (Figure 3-9). The model explained more than 99% of the variance in the group-averaged clarity ratings (Figure 3-3 E). Even better fits could be obtained for the controls by accounting for non-linearities in the effect of the increasing sensory detail on perceptual clarity beyond 16 vocoder channels, but analysis of the Akaike information criterion suggested that this increase in variance explained did not outweigh the loss of parsimony compared to the simpler model (see methods section 3.4.6 ). The simpler model was therefore retained, but all of the group differences and associations between model outputs and neurophysiology reported below remained significant if the complex model was used.
Figure 3-9 Example Bayesian model fits for a single subject in a single condition. In this case control 10 for 4 channel vocoded speech and matching written cue and vocoded word. The standard deviation of the prior and the perceptual threshold were individually optimised to best model the clarity ratings from experiment 1. Upper left: the prediction, in this case a 50% chance of the cued word (A) and an approximately zero chance of any other given word (B). Middle left: the total area under the curve of the prediction error fed back to frontal regions from temporal regions. Bottom left: the updated prediction, from which an A vs B discrimination might be performed. Upper centre: the modelled activity in a set of 21 neural units representing phonological feature categories; a weighted function of the upper right panel. Middle centre: the modelled prediction for activity in the set of 21 neural units representing phonological feature categories; a weighted function of the upper left panel. Bottom centre: The prediction error in each of the 21 modelled neural units, calculated by subtracting the prediction from the sensory input. Upper right: the probability density for the sensory identification of the cued word (A) vs any other word (B); this was defined for each subject individually as the percentage above chance in experiment 2 at the number of vocoder channels being modelled. Middle right: the area under the curve of the posterior distribution, calculated as the sum of the area under the curve of the sensory input and a weighted product of the sensory input and prior precision, with the weighting determined by the congruency of the written cue. Bottom right: the modelled clarity rating, calculated as the height of the posterior above perceptual threshold, individually normalised to a 1-4 rating scale.
Parametric and non-parametric tests indicated that patients had significantly more precise prior expectations than did controls (both p<0.01; Figure 3-3 F). There was a trend towards patients having lower perceptual thresholds (Wilcoxon p=0.08); patients required less sensory detail to give a clarity rating of 2 or higher, reflecting an appropriate downwards extension of the subjective clarity scale rather than a higher level introspective deficit resulting in patients not being ideal observers of their sensory experience (see discussion). The model results confirm that the consequence of degraded neural mechanisms for sensory predictions is not that the brain is unable to use prior knowledge (written cues) to modulate perception, but rather that patients with nfvPPA apply their prior knowledge with greater precision and inflexibility.

3.4.7 Induced oscillatory dynamics

To examine the effects of nfvPPA and task manipulations on induced oscillatory activity, I performed a time-frequency analysis of the planar gradiometer data, averaged across sensors (Figure 3-10). First, I inspected the neural instantiation of predictions by exploring induced activity during the period following presentation of the written word but before the onset of the auditory stimulus. Based on recent studies in the auditory domain I expected this updating of predictions to manifest as an increase in beta frequency oscillations preceding the onset of the spoken word (Sedley et al., 2016). This was confirmed in my cohort, with a significant increase in beta power (10 to 28 Hz) for both groups of participants beginning around 800ms after the onset of the written word, i.e. approximately 250ms before the onset of the spoken word (cluster FWE p = 0.001, Figure 3-10). At the time (992ms) and frequency (24Hz) of the peak effect for both groups overall, the single subject magnitude of the induced response correlated significantly with their precision of prior expectations as simulated by my behavioural Bayesian model (Pearson’s r(19)= -0.52, p = 0.017; Spearman’s rho = -0.54, p = 0.012). A confirmatory SPM across the whole time window confirmed the group difference implied by this relationship, with patients displaying a single cluster of greater induced beta (20 to 34Hz) power from 868ms (cluster FWE p = 0.010). There were no induced effects that were greater in controls than in patients.
Second, I complemented my evoked analysis by assessing oscillatory power during the reconciliation of predictions, i.e. after the onset of the spoken word, averaged across sensors. The results described below were not altered by subtraction of the condition-averaged evoked waveform subtracted from every trial, confirming that these are true induced responses rather than high-frequency contamination from the evoked responses described previously. Figure 3-11 A illustrates the oscillatory power induced by hearing noise vocoded speech for each group, normalised to the pre-visual stimulus baseline and averaged across the whole brain. Across all conditions, the general pattern was for increased alpha and beta power for the first ~200ms, followed by a desynchronisation from ~200ms onwards.
Figure 3-11 Induced analysis after the spoken word.

A: Total induced power after spoken word onset and main effect of cue congruency by group. B: Overall induced power difference between Match and Mismatch conditions in the alphabeta overlap range. C: Single subject time-frequency profiles for each control. The time taken to reach 80% of the peak power contrast between Match and Mismatch trials is indicated for each individual by the number below the corresponding abscissa. D: Single subject time-frequency profiles for each patient. E: Significant negative correlation between frontal grey matter volume (adjusted for age and total intracranial volume at the co-ordinates in Figure 3-7 C) and the time taken to express a congruency contrast (panel D). The grey shaded area indicates the 95% confidence band for the regression line, marked in black. F: No significant correlation between similarly adjusted superior temporal grey matter volume and effect latency.
A significant main effect of congruency was observed in both groups (Figure 3-11 A). In controls, this effect peaked at 436ms at a frequency of 12Hz (FWE p<0.001), with greater suppression of this response for spoken words that matched prior written text. In patients, a similar effect was observed, also at 12Hz, but with a later peak at 824ms (FWE p<0.001).

There were two time/frequency clusters that showed a group by congruency interaction. In the first, extending from 276-444ms between 4-24Hz (peak 300ms, 16Hz), controls displayed a greater effect of congruency than did patients (cluster FWE p<0.001). In the second, beginning at 680ms and extending beyond the end of the analysis from 6-34Hz (peak 888ms, 20Hz), the interaction was reversed, with patients showing a greater effect of congruency than did controls (cluster FWE p<0.001). The scalp distribution and source localisations of this effect are illustrated in Figure 3-12: effects were restricted to the left hemisphere and localised to areas around superior temporal gyrus. Both interaction clusters remained significant (FWE p≤0.002) in a confirmatory analysis in which power was normalised to the pre-auditory stimulus baseline. To illustrate the time-course of this oscillatory dissociation, the data from all three sensor types were restricted to 12-24Hz, encompassing the interactions in both directions, and the total effect of congruency on power in this band across the whole brain was plotted in Figure 3-11 B.
Figure 3-12 Scalp distribution and source reconstructions of induced effects.

A: Scalp topographies for the group by congruency interaction in the beta frequency band. B: Scalp topographies for the group by congruency interaction in the alpha frequency band. C: eLORETA source reconstructions for 300-450ms, 450-600ms and 600-750ms, corresponding to time windows displaying a greater effect of congruency in controls, no group by congruency interaction, and a greater effect of congruency in nfvPPA. Only left hemispheres are shown, as no significant right sided sources were demonstrated.
To investigate this delay in this beta response in individual patients, single subject time-frequency decompositions were performed and the time taken to reach 80% of the peak overall power contrast between matching and mismatching prior knowledge was defined for each subject. The effect latencies for controls were all tightly clustered between 275 and 400ms (Figure 3-11 C). Every single patient was delayed compared to every single control, with a range of 412-1048ms (Figure 3-11 D). In the patient group neural response latency was negatively correlated with grey matter density in my left frontal region of interest (r=-0.68, p=0.042; Figure 3-11 E) but not in my left temporal region of interest (r=0.34, p=0.36; Figure 3-11 F). There was a trend towards a negative relationship between latency and the standard deviation of prior expectations, though this did not reach significance (r=-0.37, p=0.1).

To summarise, all participants show the same congruency-induced reduction in activity in the STG, but nfvPPA patients are delayed in showing this response compared to controls. Thus, the differential response of patients and controls reflects a top-down effect of frontal neurodegeneration on brain responses in posterior regions that remain structurally intact (compare, Figure 3-7 B and Figure 3-12 C with Figure 3-1 E) and that respond normally to bottom-up manipulations of speech clarity (Figure 3-5, Figure 3-7 A).

3.4.8 Coherence and Connectivity during the reconciliation of predictions

To determine whether these effects are due to frontal degeneration or fronto-temporal disconnection, I examined coherence and connectivity between the frontal and temporal lobe sources of interest (Figure 3-7 C) during the 900ms immediately following the onset of each spoken word. I employed two MEG connectivity analysis methods that give complementary information concerning fronto-temporal dynamics during the reconciliation of predictions. First, Imaginary Coherence, which is immune to volume conduction effects and source spread (Nolte et al., 2004) as well as differences in power (Bowyer, 2016), allowing us to be confident that relationships I describe are true reflections of the underlying brain dynamics. Both groups had significant fronto-temporal coherence up to around 25 Hz (Figure 3-13 A). Coherence in the beta band (13-23Hz) was significantly stronger in patients than controls (Figure
This suggests that an overall reduction of fronto-temporal connectivity cannot explain my observed differences between nfvPPA patients and controls.

Figure 3-13 Coherence and connectivity analysis.
For the time series of frontal and temporal sources of interest between 0 and 900ms after every spoken word onset. The evoked waveform was subtracted from the time series of frontal and temporal source activity between 0 and 900ms after every spoken word onset before analysis. Horizontal lines at the top of each plot denote frequencies at which the line of matching colour statistically exceeds either the null distribution (A and C), its counterpart group (B) or differs from zero (D). A: Imaginary coherence. The median of the observed inter-source coherence is shown in black. 1000 randomisations of the null distribution are shown in grey. B: Imaginary coherence by group. Shading represents standard error of the mean. C: Granger causality. Median influences from temporal to frontal sources are shown in green and frontal to temporal sources in blue. 1000 randomisations of the null distribution are shown in grey. D: Relative normalised Granger causal relationships between temporal and frontal sources by frequency. Grey shading represents the standard error of the directionality contrast in a repeated measures general linear model.
Imaginary coherence does not provide robust indices of directionality, because inter-regional interactions potentially occur over more than one oscillatory cycle. I therefore also examined Granger causal relationships between my sources of interest (Granger, 1969, Gow et al., 2008). This metric allows us to look at the directionality of non-zero-lag fronto-temporal interactions while still being relatively robust to volume conduction. Grainger causality tests whether information from the past activity of one region can predict future activity in another better than its own past. Highly significant bi-directional Granger causal relationships were observed between temporal and frontal sources (Figure 3-13 C). To compare these, while avoiding confounds due to differences in signal to noise ratio between regions and between individuals (which can alter the magnitude of Granger Causal relationships (Nolte et al., 2004)), I divided the magnitude of each frequency value by the across-frequency mean for each individual and region to create a profile of relative influence for each region at each frequency. This demonstrated significantly stronger temporal to frontal Granger Causal relationships at low frequencies, while frontal to temporal influences were stronger at higher, beta band, frequencies (Figure 3-13 D). These findings are in agreement with a recent study of MEG connectivity during written language comprehension (Schoffelen et al., 2017).

Overall, my finding of increased imaginary coherence in the patient group in a frequency band where frontal to temporal Granger causal influences predominate demonstrates that frontal neurodegeneration has increased rather than reduced top-down connectivity from frontal to temporal regions.

3.5 Discussion:

The principal finding of this study is that neurodegeneration of the frontal language network results in the delayed neural resolution of predictions in temporal lobe. Conversely, the temporal lobe neural responses to bottom-up manipulations of sensory detail were not delayed. This is evidence that the resolution of sensory predictions is mediated by frontal regions in humans. My source-space analysis demonstrates that frontal regions are working harder overall in nfvPPA patients. My demonstration of increased coherence in the patient group, in a frequency band where frontal to
temporal influences predominate, is consistent with degraded neural mechanisms in frontal regions requiring increased fronto-temporal interaction to reconcile excessively precise predictions. This view is supported by recent observations that left inferior frontal imaginary coherence is decreased in nfvPPA during the resting state (Ranasinghe et al., 2017), confirming that the increased fronto-temporal coherence I observe here reflects specific engagement of these mechanisms during language perception and not simply a global upregulation. Together, these findings resolve a key controversy in speech perception, by demonstrating that frontal regions are playing a core role in reconciling predictions with sensory input during the perception of speech (Abdelnour et al., 2014, Park et al., 2015, Raj et al., 2015).

This impairment of predictive processing has significant perceptual consequences. Most strikingly, frontal neurodegeneration does not reduce the degree to which the brain employs contextual prior knowledge to guide lower-level speech perception. Rather, through Bayesian modelling, I show that prior knowledge of expected speech content is applied in an overly precise or inflexible fashion, thereby producing a larger-than-normal behavioural effect of prior knowledge on nfvPPA patients’ ratings of speech clarity. This is not an effect specific to my experimental manipulation; it can explain why patients with nfvPPA complain of difficulties with understanding speech, out of keeping with their auditory sensitivity (see section 5.2.1), why they perform poorly on tasks requiring the comparison of two auditory stimuli (Goll et al., 2010, Grube et al., 2016) (section 5.2.2) despite being able to perform other forced choice tasks well, and why they display receptive agrammatism (section 5.2.3). In validation of computational and theoretical models of predictive coding (Arnal and Giraud, 2012, Bastos et al., 2012, Sedley et al., 2016), I demonstrate that the precision of subjects’ predictions correlates with the magnitude of induced beta-frequency oscillations, which have recently been shown to correspond temporally and quantitatively to the updating of predictions (Sedley et al., 2016).

Previous neuro-imaging evidence has suggested that frontally-mediated top-down predictions during speech perception are able to explain MEG response magnitudes (Sohoglu et al., 2012, Sohoglu and Davis, 2016) and fMRI pattern information in the superior temporal gyrus (Blank and Davis, 2016). Here, I go beyond these associations and provide causal evidence for a functional contribution of frontal networks in
supporting top-down predictions by demonstrating that disruption of these networks has the remote effect of delaying the reconciliation of predictions in temporal lobe regions that are anatomically intact, with striking behavioural consequences.

The behavioural data, Bayesian modelling and neurophysiological results all support the proposal that perceptual predictions operate within an hierarchical generative network, which for speech perceptions spans auditory, superior temporal, and inferior frontal cortices.

3.5.1 Fronto-temporal dissociations in nfvPPA

Patients with nfvPPA lacked normal modulation of frontal neural activity by cue congruency, were delayed in engaging frontal regions, and showed greater frontal activity than control participants overall. This is analogous to the observation that elderly controls globally upregulate cognitive control networks that are selectively engaged by younger listeners only when speech is degraded, and thus appear to lack difficulty-related modulation of activity (Erb and Obleser, 2013). My work goes significantly beyond these previous findings, however, in showing that this frontal deficit also affects neural responses in (anatomically unaffected) posterior regions.

In contrast to these impairments, patients displayed normal power of evoked neural responses for analyses combined over sensors and had normal neural responses to changes in sensory detail shown here and in previous work to localise to temporal regions (Sohoglu et al., 2012) (Figure 3-7 A). These observations provide reassurance that the lack of atrophy in auditory regions of temporal lobe does not mask a microscopic abnormality in auditory neural function. Similarly, it is reassuring that in both groups clarity ratings were similar for mismatching and neutral cues, and the response to questionnaires suggested similar ecological perception of speech in most conditions. This excludes trivial explanations for my behavioural results such as patients becoming confused by mismatches or having an altered approach to subjective rating scales. The group difference in uncued identification of vocoded speech was small, and was accounted for in my Bayesian modelling, by using the results of experiment 2 to individually define the precision of the sensory input for each subject and number of noise vocoder channels. Thus, my observation of abnormal effects of
cue congruency in nfvPPA cannot easily be explained by basic auditory processing deficits, or by higher signal detection thresholds.

Most strikingly, patients displayed a significant delay in the effects of cue congruency in temporal lobe in my sensor space, source space and induced analyses. Under the predictive coding framework, these effects arise because the integration of prediction error is an iterative process, whereby predictions are recursively updated in the light of sensory input to minimise error (Blank and Davis, 2016). For the patients with nfvPPA, the degraded neural architecture and aberrantly precise predictions might mean that this updating requires more iterations and/or that each iteration takes more time. As a consequence, more neural activity is observed and reconciliation of predictions with sensory signals is delayed.

3.5.2 Frontotemporal hierarchy and predictive coding

My experiments have demonstrated that frontal lobes have top-down causal influences on neural activity in temporal lobe. This requires hierarchical sensory processing, with feed-back and feed-forward of information between levels. Predictive coding is one framework to understand such processing, and makes a number of specific claims about the nature of top-down and bottom-up signals. Here I discuss the evidence for predictive coding, and the inability of alternative frameworks to explain my findings.

Firstly, the spatial and temporal pattern of neurophysiological responses in elderly control participants replicated those previously demonstrated in young people (Sohoglu et al., 2012), which have been successfully modelled in a predictive coding framework (Sohoglu and Davis, 2016), and which cannot be effectively modelled by sharpening theories of neural representation (Blank and Davis, 2016). Current instantiations of predictive coding models state that each stage of the processing hierarchy passes forwards to the next stage only prediction errors and backwards only predictions (Hinton, 2007, Friston, 2008). An alternative hypothesis – which might explain some of the present data – is that prior expectations are set up in frontal regions but not fed back to superior temporal regions (Norris et al., 2000, Norris et al., 2016). When auditory input is received, a process of sensory analysis begins in temporal lobe, with the output fed forwards to frontal regions in real-time, where a matching process
occurs. If frontal regions detect that the auditory information matches expectations, they indicate that further processing is unnecessary by feeding back a stop signal to temporal regions. Such a mechanism could also account for my observation of a delayed reduction in superior temporal activity when the text cue is matching. This alternative hypothesis would also predict that total superior temporal lobe activity over the whole epoch would be greater in patients than in controls, as their stop signal is delayed. The present data showed a non-significant trend in this direction. However, this stop-signal hypothesis is unable to account for fMRI evidence demonstrating an interaction of prior knowledge and sensory detail in superior temporal representations of degraded speech (Blank and Davis, 2016). These fMRI findings can only by simulated by a computational model in which superior temporal regions represent the discrepancy between predicted and heard speech (i.e. prediction error). In the predictive coding model, the delayed neural effects of cue congruency observed here reflect an iterative process whereby predictions are recursively updated to minimise error; this process operates more slowly in my patient group.

Most models of prediction and perception, including my own Bayesian modelling, make the assumption that perceptual outcomes represent an ideal observation of peripheral sensation. This might not be the case if individuals hold aberrant beliefs about the fidelity of their sensory input based on differences in previous experience (Orhan and Jacobs, 2014). The results of my Bayesian modelling are inconsistent with the view that my patients are not ideal observers of their sensory experience. If patients with nfvPPA had learnt that their auditory input were unreliable, this could only explain the present data by proposing a dissociation between an underestimation of the precision of their sensory input when reporting perceptual clarity (experiment 1) and an intact ability to discriminate sensory features when distinguishing alternative vocoded words (experiment 2) (Norris et al., 2016). This would manifest in my Bayesian modelling as an increase in perceptual threshold, as any given distribution of sensory input would be reported as less clear. In fact, I found that patients did not statistically differ from controls in terms of their perceptual thresholds. Indeed, the trend was towards lower thresholds, which might reflect an appropriate downwards extension of the bottom-end of their perceptual clarity rating scale to reflect the fact that they were slightly less good at identifying vocoded speech than controls.
(experiment 2), indicating that they had access to slightly less sensory detail during the experiment. Therefore, my behavioural results cannot be accounted for by patients with nfvPPA not being ideal observers of vocoded speech.

The hypothesis that beta oscillations represent the instantiation of predictions has existed for some time (see Arnal and Giraud, 2012 for review). It has been supported by evidence including computational simulations (Bastos et al., 2012), and empirical observations of backward beta connectivity in speech processing (Fontolan et al., 2014). More recently, direct recordings from human auditory cortex have directly linked beta frequency oscillations to the updating of predictions (Sedley et al., 2016), based on correlations between observed brain activity and Bayes-optimal predictions generated from presented stimuli. In my cohort, I not only replicate the finding of beta oscillations as a correlate of prediction instantiation, but go further in demonstrating that, irrespective of disease status, the strength of this beta activity across subjects relates to the precision of their predictions, as determined by my Bayesian behavioural modelling.

3.5.3 Study limitations

This experiment relies on a dissociation in neurodegeneration between affected frontal brain regions and unaffected superior temporal regions. In section 3.4.1 I have demonstrated this to be true in terms of grey matter volume, however neurodegeneration caused by tau and TDP-43 pathology manifests in synaptic dysfunction before volume loss becomes evident. While I cannot exclude such early changes in my cohort, in section 3.4.3 I demonstrate that measured physiological responses to bottom-up manipulations of sensory detail are normal, and that it is only through the manipulation of prior expectations that a group difference in the timecourse of neuronal is evident.

In section 3.4.5 I demonstrated that patients with nfvPPA do not perform as well as controls at identifying vocoded speech. It is therefore possible that an overall difference in task difficulty affected the degree to which participants relied on priming information. However, the group performance difference was small, and was accounted for on an individual basis in the Bayesian modelling of section 3.4.6.
Additionally, patients performed better at identifying speech with 8 channels than controls did with 4 channels, but still displayed a much larger congruency effect for 8-channel vocoded words than controls do for 4-channel speech.

Finally, as nfvPPA is a very rare disease and only early-stage patients were assessed, the number of individuals in this study is relatively small. The principal finding of a neuronal processing delay dependent on prior expectations was manifest in every patient individually (Figure 3-11), and is therefore not in doubt. However, while statistically significant, the correlation analyses between induced beta power and prediction strength (Figure 3-10), and between processing latency and frontal grey matter volume (Figure 3-11) would be strengthened by replication.

3.6 Conclusion:

I demonstrate distant neural effects of the degeneration of top-down signals from frontal lobes, during speech perception. Patients with neural damage in these frontal regions displayed significantly delayed activity in temporal regions that I have demonstrated to function normally in response to bottom-up manipulations of speech clarity. This did not result in an inability to utilise contextual information, but was rather associated with inflexible and overly precise prior expectations. I provide evidence of a direct relationship between the degree of frontal lobe degeneration and a delay in the neural mechanism for the reconciliation of predictions, which results in their inflexible application. In turn, this results in the formation of aberrantly precise prior expectations, perhaps in order to maximise their benefit in the more challenging states of a dynamic environment. These aberrantly precise priors manifest as increased beta power during the instantiation of predictions, in agreement with theoretical frameworks of predictive coding. Together, my results provide causal evidence for a critical role of frontal regions for the reconciliation of predictions during the perception and comprehension of speech.
3.7 Chapter Summary:

Perception relies on the integration of sensory information and prior expectations. Studies of speech perception provide evidence for left-lateralised top-down influences of frontal activity on temporal brain regions, which might support hierarchical generative models of sensory input such as predictive coding. However, the role of top-down predictions in supporting higher level perception remains controversial. Here I show that selective neurodegeneration of human frontal speech regions results in significantly delayed reconciliations of predictions in temporal cortex. These temporal regions were not atrophic, displayed normal evoked magnetic and electrical power, and preserved neural sensitivity to (bottom-up) manipulations of sensory detail. Frontal neurodegeneration did not prevent the use of contextual information: instead, prior expectations were applied inflexibly. The precision of predictions correlated with beta power during the instantiation of predictions, in line with theoretical models of the neural instantiation of predictive coding. My results provide evidence that fronto-temporal interactions play a causal role in reconciling prior predictions with degraded sensory signals, thereby demonstrating that top-down frontal mechanisms are necessary for effective use of prior knowledge during the perception of speech. This work demonstrates that higher level frontal mechanisms for cognitive and behavioural flexibility make a critical functional contribution to the hierarchical generative models underlying speech perception.
Chapter 4: Artificial grammar learning in vascular and progressive non-fluent aphasias

4.1 Preface

Chapter 3 provided a link between the network-level physiological changes that occur in neurodegeneration and the deficits they cause in core cognitive systems, by providing evidence for diaschisis; abnormalities in top-down influences from frontal brain regions caused changes in neural processing in temporal brain regions that were functionally and anatomically intact in isolation. In this chapter I further explore deficits in core cognitive systems by using psychophysical methods to investigate the nature of grammatical impairment in frontal aphasia. To do this I used an artificial grammar learning paradigm to define the abilities of patients with agrammatism due to nfvPPA, in comparison to controls and patients with agrammatic aphasia due to frontal stroke. This chapter concentrates specifically on the nature of grammatical processing as a core cognitive system, but lays the groundwork for chapter 5, section 5.2.3 in which I integrate the results of chapters 3 and 4 to directly relate abnormal neurophysiological connections in the language network to the complex behavioural abnormalities observed in nfvPPA.

This chapter is based on a paper I have recently published in Neuropsychologia (Cope et al., 2017b). It represents a multi-site study, requiring the help of many hands. I designed and piloted the artificial grammar in conjunction with Dr Benjamin Wilson and Prof Christopher Petkov, as it flows naturally from their studies of the grammatical abilities of non-human primates. I collected the data for all of the patients with nfvPPA, all of the controls, and one of the stroke patients; the rest of the stroke patients were tested in Reading by Ms Rebecca Drinkall under supervision of Dr Holly Robson, who followed my protocol and used computer programmes written by me for stimulus delivery. Stereotyped speech samples were recorded for each participant, and were triple-rated according to the Boston Diagnostic Aphasia Examination scales by myself, Dr Holly Robson and Prof Karalyn Patterson. Ms Julie Wiggins (research nurse) very kindly helped me to obtain structural MRI scans for the nfvPPA group. I
performed the voxel based morphometry analysis of these scans with assistance from Mr P Simon Jones. Lesion mapping of the stroke group was performed by Dr Lauren Dean (because many had only clinical MRI scans unsuitable for voxel based morphometry). I am grateful to Prof Timothy Griffiths (Newcastle) for initial discussions about study design, and for allowing me to recruit one of his patients with nfvPPA, one of his patients with frontal stroke aphasia, and several of the neurological controls to the final study. I am also indebted to Professor Matthew Lambon-Ralph (Manchester) for allowing me access to some of his well characterised patients with aphasia due to frontal stroke for the purposes of sequence development and piloting (these pilots do not form part of the final study reported here, but were invaluable in ensuring that the stimuli and paradigms would be accessible to my clinical cohorts). I performed all of the analyses, made all of the figures except the lower half of panel B of Figure 4-1 (which was made by Dr Lauren Dean) and Figure 4-4 (which was made by Dr Benjamin Wilson). I wrote the manuscript on which this chapter is based with the aid of insightful comments from the collaborators listed above, especially Dr Wilson and Prof Petkov, as well as Dr Manon Grube and my supervisor, Prof James Rowe.
4.2 Introduction

Aphasia is an impairment of speech and language that often leaves other cognitive and intellectual capacities preserved. Patients with non-fluent aphasias due to frontal lobe damage exhibit significant impairments in grammar (Caramazza and Zurif, 1976, Caplan et al., 1985, Berndt et al., 1996). The grammatical impairments in comprehension and production are separable, but tend to be highly correlated (Berndt et al., 1983), suggesting that they stem from disruption of core syntactic processes rather than processes such as memory, executive function or motor function (Wilson et al., 2011). The deficits are phenomenologically similar in patients with damage due to neurodegeneration (non-fluent variant Primary Progressive Aphasia, nfvPPA) and stroke (‘Broca’s aphasia’); however detailed analysis of speech output has revealed somewhat differential impairments (Patterson et al., 2006, Thompson et al., 2013). Impairments of receptive abilities have not previously been compared in similar detail.

Beyond these linguistic deficits, patients with aphasia also display auditory domain general processing deficits that are not specifically related to language (Caramazza and Zurif, 1976, Dominey et al., 2003, Patel et al., 2008, Christiansen et al., 2010, Goll et al., 2010, Grube et al., 2012, Geranmayeh et al., 2014a, Zimmerer et al., 2014a, Zimmerer and Varley, 2015, Grube et al., 2016). Such studies have raised the possibility that deficits in structured sound processing may play a prominent role in language disorders, but the nature and extent of these deficits remain unclear. It also remains unclear whether impairments in aphasia are specific to the speech domain (Conway and Pisoni, 2008), or also apply to non-linguistic auditory sequences (Christiansen et al., 2010). One study identified impairments in implicit musical sequence learning in vascular aphasia (Patel et al., 2008), but direct comparisons outside of a musical framework are lacking. If artificial grammar learning tasks tap into domain general (rather than language specific) processes, one would expect rule acquisition to generalise from sequences of nonsense words to identically structured sequences of other sounds, such as tones.

It has been commonly held that grammatical impairments are specific to complex linguistic constructs such as hierarchical relationships and the passive voice (Goodman and Bates, 1997, Grodzinsky, 2000), but there is limited evidence for such dissociations.
(Zimmerer et al., 2014b). By contrast, some studies suggest that the processing of adjacent relationships may be disproportionately impaired by frontal lesions involving motor association cortex (Opitz and Kotz, 2012). Recent studies examining artificial grammar learning in agrammatic aphasia secondary to stroke have focussed on linear sentential structures with varying transitional probabilities (Schuchard and Thompson, 2017). A key outstanding question, therefore, is whether agrammatic aphasia is characterised specifically by deficits for more complex linguistic structures or rather by a more global impairment in processing structured auditory sequences (Berndt, 2000).

Artificial grammar learning tasks are particularly well suited for delineating competence in structured sequence processing, as they focus on ordering relationships in the absence of other cues (e.g., semantics, phonology or pragmatics). They test learning of the rules governing the order in which stimuli occur in a sequence (Reber, 1967). Participants are typically exposed to sequences of stimuli that follow certain rules, so that the ordering relationships between the sequence elements can be learned implicitly. They are then tested with novel sequences that are either consistent with these rules or that violate them in some way, to assess learning. The implicit nature of these tasks allows the testing of a wide range of participants, including patients with aphasia. Unlike natural language tasks, it is possible to present structurally identical sequences comprised of different tokens, for example nonsense words or non-linguistic tone stimuli, to assess the contribution of phonological processing. Finally, artificial grammars with multiple levels of complexity can be used to quantify how well participants are able to learn increasingly complex rules, which may more closely reflect those in natural language grammars (Romberg and Saffran, 2013, Wilson et al., 2015).

The ability to process auditory sequences, even when stimuli are meaningless, is strongly linked with linguistic proficiency (Gómez and Gerken, 2000, Conway and Pisoni, 2008, Conway et al., 2010, Frost et al., 2015). Neuroimaging studies have demonstrated that artificial grammar processing engages a left-lateralised network of frontal, temporal and parietal brain areas similar to the set of regions involved in syntactic operations during natural language tasks (Friederici et al., 2000, Ni et al., 2000, Friederici and Kotz, 2003, Petersson et al., 2004, Forkstam et al., 2006, Friederici et al., 2006, Hickok and Poeppel, 2007, Bahlmann et al., 2008, Makuuchi et
Successive processing ability of patients with non-fluent aphasia patient groups has not been systematically compared across aetiologies. Non-fluent variant Primary Progressive Aphasia (nfvPPA) is in many ways the neurodegenerative equivalent of Broca’s aphasia, being an adult onset neurodegenerative aphasia characterised by agrammatism and speech apraxia (Gorno-Tempini et al., 2011), although some differences do exist in the pattern of speech output impairment (Patterson et al., 2006). The majority of cases are associated with primary tau pathology but a significant minority have TDP-43 related disease (Kertesz et al., 2005, Josephs et al., 2006, Knibb et al., 2006ab, Knibb et al., 2006ba, Mesulam et al., 2014). nfvPPA typically leads to subtle neuroimaging changes in left inferior frontal and insular cortex (Gorno-Tempini et al., 2004), which correlate with clinical severity (Rogalski et al., 2011). Chronic non-fluent aphasia due to stroke (Broca’s aphasia) results in a similar clinical phenotype of agrammatism and apraxia of speech. The left frontal tissue damage is stable, with partial clinical improvement over time (Kertesz and McCabe, 1977). The extent and pace of this improvement is variable and depends strongly on the integrity of the underlying white matter (Price et al., 2010, Seghier et al., 2016). Better understanding of the abilities of participants with similar symptoms arising from very different aetiologies could provide valuable insights into the neurobiological underpinnings of domain-general and language-related processes, and inform treatment strategies (Brownsett et al., 2014, Geranmayeh et al., 2014a).

In the present study, patients with nfvPPA, non-fluent aphasia due to stroke, and matched controls were tested on their implicit learning of a mixed-complexity artificial grammar, combining sequencing relationships of increasing complexity using nonsense words or tones. I aimed to test the following linked hypotheses:

1) Rule acquisition differs when structurally identical sequences are comprised of nonsense words rather than non-linguistic tones.

2) Artificial grammar learning ability is similar in patients with vascular and neurodegenerative aphasia.
3) Grammatical impairments in aphasic patients are disproportionately greater for complex, configurational or hierarchical, sequencing operations.

4) Patients with aphasia can improve their ability to detect grammatical disruptions with repeated implicit training.

### 4.3 Methods

#### 4.3.1 Participants

Three groups of participants were recruited. The nfvPPA group overlapped with, but was not identical to, that recruited for the work described in chapter 3. Demographics of the groups are outlined in Table 4-1.

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>nfvPPA</th>
<th>Stroke</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>11</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Age</td>
<td>69 (8, 54-79)</td>
<td>73 (7, 63-82)</td>
<td>60 (11, 33-74)</td>
</tr>
<tr>
<td>Age leaving education</td>
<td>18 (2, 15-22)</td>
<td>18 (3, 15-25)</td>
<td>20 (4, 15-26)</td>
</tr>
<tr>
<td>Years of musical training</td>
<td>2 (3, 0-10)</td>
<td>1 (1, 0-3)</td>
<td>3 (5, 0-13)</td>
</tr>
</tbody>
</table>

**Table 4-1 Participant demographics**

*Mean (s.d., range). Age leaving education is reported as it is a better measure of highest scholastic attainment than number of years in study. No individuals were mature students.*

All patients were right handed. One control was left handed. Thirteen patients with mild to moderate nfvPPA were identified from specialist cognitive clinics according to consensus diagnostic criteria (Gorno-Tempini *et al.*, 2011). These criteria were strictly applied; particular care was taken to exclude non-fluent patients who had lexical difficulties, in order to select patients most likely to have underlying tau or TDP-43 related pathology preferentially involving left frontal lobes, rather than Alzheimer-type pathology of parietal lobes (Rogalski *et al.*, 2011, Rohrer *et al.*, 2012, Sajjadi *et al.*, 2012, Sajjadi *et al.*, 2014, Mandelli *et al.*, 2016). Three patients were excluded on the basis of yes/no response confusion (a common early symptom in nfvPPA that might otherwise have reduced my power to detect language specific effects), resulting in 10
complete nfvPPA datasets. On the short form of the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass et al., 1983) all patients scored 10/10 for responsive naming, 12/12 for special categories, at least 15/16 for basic word discrimination and at least 9/10 for following complex commands. While nfvPPA exists on a spectrum, with differing ratios of speech apraxia and agrammatism, all of my patients displayed some degree of impairment of expressive grammar in free speech, and all but two displayed impairment of receptive grammar as measured by the sentence comprehension task on the ‘verb and sentences test’ (VAST) (Bastiaanse et al., 2003) (mean 87.5%, range 70-100%). Similarly, the patients varied in their degree of expressive agrammatism, but none was completely unimpaired. Speech profiles are shown in Figure 4-1 A. I determined BDAE profiles as the average of three independent ratings of free speech and picture description undertaken by Dr Holly Robson, Prof Karalyn Patterson, and myself. Inter-rater reliability for grammatical form was good, with pairwise Pearson correlations of 0.87, 0.85 and 0.85. Areas of significant grey or white matter loss are shown in Figure 4-1 B, upper panel.
Figure 4-1 Group characteristics
A) Boston Diagnostic Aphasia Examination Profiles for nfvPPA and stroke groups. Normal values illustrated as a broken black line. Colour coding of individual profiles based on Aphasia Severity Rating Scale; 1 = red, 2 = magenta, 3 = yellow, 4 = blue. No patients had an ASRS of 0 (no usable speech or auditory comprehension) or 5 (minimal discernible handicap). B) Upper: Voxel based morphometry of nfvPPA vs age-matched healthy controls. Coloured regions demonstrate cluster-wise significance at FWE<0.05 with a cluster defining threshold of 0.001 for either grey or white matter volume. Lower: lesion overlap map for the stroke group.
It is widely recognised that patients with nfvPPA report difficulties with understanding speech (Goll et al., 2010, Cope et al., 2014b, Grube et al., 2016). I asked the patients in this study to complete Likert scales assessing their difficulty with ‘Understanding speech in a quiet room’, ‘Telling the direction a sound is coming from’, ‘Understanding speech in a noisy restaurant’, ‘Hearing announcements at a bus or rail station’ as well as ‘How loud do people tell you your TV is?’ Compared to a matched group of controls, patients differed only in reporting more difficulty with ‘Understanding speech in a quiet room’ (p=0.02). Of the Boston Diagnostic Aphasia Examination (BDAE) sub-scores, this difficulty was strongly correlated only with ‘Grammatical form’ ($r^2=0.778$, p<0.001, Figure 4-2).

Figure 4-2 A plot for nfvPPA of difficulty ‘Understanding speech in a quiet room’ against ‘Grammatical form’ from the BDAE. Individual patients are shown as blue circles. A boxplot of control ratings is shown in red. The dashed trend line had an $r^2$ of 0.778 (p<0.001).

Patients with non-fluent aphasia due to left sided stroke were recruited from a volunteer database, supplemented by the identification of incident cases by regional research networks. Recruitment criteria were: a single stroke of at least six months chronicity resulting in at least one month of non-fluent aphasia, with MRI evidence of
involvement of either left inferior trigone or operculum. Samples of speech from the participants are available in supplementary materials, and speech profiles (again triple marked) are shown in Figure 4-1 A. On the whole, the stroke group had more severe language impairments than the nfvPPA group. All had some degree of impairment of grammatical form in free speech, and all but two had impairment of receptive grammar on the VAST (mean 70%, range 40-100%). Lesion overlap maps are shown in Figure 4-2 B, lower panel.

Care was taken to recruit an appropriate control group. During development of the artificial grammar, extensive piloting developed structures for which learning was least influenced by years of education or performance on global cognitive tests. Nonetheless, it is important to minimise this potential confound by avoiding the use of biased volunteer panels, which tend to preferentially recruit highly educated individuals with supra-normal motivation in research tasks. Therefore, I recruited 8 neurological controls with either chronic inflammatory demyelinating polyneuropathy or multifocal motor neuropathy with conduction block, and three spouses of patients with nfvPPA, resulting in 11 control datasets. These individuals were chosen to represent a cohort of age-matched individuals with healthy brains and similar levels of habitual neurological contact to the patient groups. All scored normally on the Addenbrooke’s Cognitive Examination – Revised (ACE-R) (mean 96/100, range 92-99) and Raven’s progressive matrices (mean 47/60, range 37-60).

It was possible to perform pure tone audiometry in all patients with nfvPPA, 9 of the 12 patients with stroke aphasia and 8 of the 11 neurological controls. This demonstrated that the groups had well matched and age-appropriate auditory acuity (Figure 4-3).
Figure 4-3 Composite pure tone audiograms for each group. Red circles joined by lines denote the median auditory threshold in the right ear; blue crosses joined by lines the median auditory threshold in the left ear. Shaded areas represent the inter-quartile range of auditory acuity in the right (red) and left (blue) ears.

4.3.2 Stimuli

The Artificial Grammar (AG) used here generates sequences of stimuli from 8 unique elements (Figure 4-4). These sequences are governed by a number of rules of increasing complexity. Rule 1) if a ‘C’ element occurs it must be immediately followed by a ‘D’ element. This represents a simple, invariant linear relationship between two adjacent sequence elements, and will henceforth be referred to as the ‘linear’ rule. Rule 2) all of the ‘A’ elements in the sequence must occur before all of the ‘B’ elements. This is a more complex rule, requiring the participants to recognise a general property of the sequences, and will henceforth be referred to as the ‘configurational’ rule (Zimmerer and Varley, 2015). Rule 3) each ‘A’ element type must be paired with the appropriate ‘B’ elements in embedded relationships (e.g., $A_1[A_2[A_3B_3]B_2]B_1$). This complex operation requires tracking both the number and the order of the ‘A’ elements and matching these to the subsequent ‘B’ elements, and is referred to as the ‘hierarchical’ rule.
Sequences consistent with the AG are generated by following any path of arrows from start to end in the illustrated state transition graph (Figure 4-4). Ten consistent sequences were used for the exposure phase of the experiment (Table 4-2). These sequences were of variable length and contained all of the legal transitions possible with the AG. The remaining subset of consistent sequences generated by the AG were kept for the subsequent testing phase, to allow us to present novel, previously unheard sequences (Table 4-2). During the testing phase, the participants were presented with 4 repetitions each of 6 consistent and 6 violation sequences, in a pseudo-random order. All test sequences were six elements long, meaning that sequence duration could not be used as a cue by participants. The six violation sequences contained two sequences with violations of each of the three AG rules (Table 4-2). This design allowed us to identify the specific features of the sequences to which the participants were sensitive.
I tested participants with identically structured sequences of both naturally spoken consonant-vowel-consonant (CVC) nonsense words and non-linguistic tone stimuli. The stimuli were designed to provide acoustic cues to highlight the relationships between some of the key sequencing relationships, as follows. In both the CVC and tone experiments, the ‘A’ and ‘B’ elements fell into distinct acoustic categories. In the CVC experiment the ‘A’ elements all took the form “s-vowel-f” (e.g., “sif”) while the ‘B’ elements were “g-vowel-k” (e.g., “gik”). In the tone experiment the ‘A’ elements were all upwards pitch sweeps while the ‘B’ elements were downward sweeps. Furthermore, the A\textsubscript{X}-B\textsubscript{X} relationships that are critical to Rule 3 were highlighted by the presence of the same vowel sounds in the nonsense word experiment (i.e., A\textsubscript{1} and B\textsubscript{1} both contain the central vowel ‘a’) or the tone height in the tone experiment (i.e., both the A\textsubscript{1} and B\textsubscript{1} pitch sweeps are centred on the same frequency). To ensure that the participants learned aspects of the AG during the exposure phase, rather than simply responding to acoustical properties of the stimuli, I designed the tone stimuli to avoid linearly increasing or decreasing in pitch in the A\textsubscript{1}A\textsubscript{2}A\textsubscript{3} or B\textsubscript{3}B\textsubscript{2}B\textsubscript{1} parts of the sequences. Instead, the centre frequencies of the tone sweeps in such a sequence would be ‘mid-low-high-high-low-mid’. The ‘C’ and ‘D’ elements in the nonsense word experiment were designed to be clearly phonetically distinct from the ‘A’ and ‘B’ stimuli, and in the tone experiments they were continuous pure tones of high or low pitch.

<table>
<thead>
<tr>
<th>Exposure Sequences</th>
<th>Testing Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>A\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1}</td>
<td>A\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1} Consistent Familiar</td>
</tr>
<tr>
<td>CDA\textsubscript{1}B\textsubscript{1}CD</td>
<td>A\textsubscript{1}B\textsubscript{1}CDCD Consistent Familiar</td>
</tr>
<tr>
<td>A\textsubscript{1}B\textsubscript{1}CDCDC</td>
<td>CDA\textsubscript{1}B\textsubscript{1}CD Consistent Familiar</td>
</tr>
<tr>
<td>A\textsubscript{1}A\textsubscript{2}B\textsubscript{2}B\textsubscript{1}</td>
<td>A\textsubscript{1}A\textsubscript{2}B\textsubscript{2}B\textsubscript{1} Consistent Novel</td>
</tr>
<tr>
<td>CDA\textsubscript{1}B\textsubscript{1}</td>
<td>CD CDCDA\textsubscript{1}B\textsubscript{1} Consistent Novel</td>
</tr>
<tr>
<td>A\textsubscript{1}B\textsubscript{1}CD</td>
<td>CDCDA\textsubscript{1}B\textsubscript{1} Consistent Novel</td>
</tr>
<tr>
<td>CDCDA\textsubscript{1}B\textsubscript{1}CD</td>
<td>CDCDA\textsubscript{1}B\textsubscript{1} Consistent Novel</td>
</tr>
<tr>
<td>CDA\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1}</td>
<td>CDCDA\textsubscript{1}B\textsubscript{1} Consistent Novel</td>
</tr>
<tr>
<td>A\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1}CD</td>
<td>DCA\textsubscript{1}B\textsubscript{1}DC Violation Violates Rule 1</td>
</tr>
<tr>
<td>A\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1}</td>
<td>CDCB\textsubscript{1}A\textsubscript{1}CD Violation Violates Rule 2</td>
</tr>
<tr>
<td>A\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1}</td>
<td>A\textsubscript{1}B\textsubscript{1}A\textsubscript{2}B\textsubscript{2}CD Violation Violates Rule 1</td>
</tr>
<tr>
<td>A\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1}</td>
<td>A\textsubscript{1}A\textsubscript{2}B\textsubscript{1}B\textsubscript{2}CD Violation Violates Rule 2</td>
</tr>
<tr>
<td>A\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1}</td>
<td>A\textsubscript{1}A\textsubscript{2}A\textsubscript{3}B\textsubscript{3}B\textsubscript{2}B\textsubscript{1} Violation Violates Rule 3</td>
</tr>
</tbody>
</table>

Table 4-2 Exposure and testing sequences.
The nonsense words were produced by a female speaker, recorded with an Edirol R-09HR (Roland Corp.) sound recorder, and combined into exposure and testing sequences using Matlab (100ms inter-stimulus intervals, ISI). The average duration of the nonsense words was 477ms (standard deviation = 7ms). The tone stimuli were generated using Matlab. The low tone sweeps were linear sweeps between 100 and 150Hz (i.e., A2 began at 100Hz and increased to 150Hz, B2 began at 150Hz and decreased to 100Hz). The middle tone sweeps ranged between 200 and 300Hz and the high tone sweeps ranged between 400 and 600Hz. The C and D stimuli were pure tones at 350 and 800Hz respectively. The duration of all tones was 450ms, and these were combined into sequences with ISIs of 100ms. Stimuli were presented through Sennheiser HD250 linear 2 headphones, driven by either an Edirol UA-4X or Behringer UCA 202 external sound card. The amplitudes of all stimuli were root-mean-square (RMS) balanced, and sequences were initially presented to participants at ~75 dB SPL (calibrated with an XL2 sound level meter, NTI Audio). At the start of the exposure phase, participants were asked if this volume was comfortable and clearly audible and, if not, were allowed to freely adjust the volume to their preference.

Exposure sequences and test sequences for the CVC and tone languages were presented in exactly the same fixed pseudo-random order. There were no differences between the orders of sequences in the CVC and tone runs; only the sound tokens used to represent each element in the artificial grammar differed.

### 4.3.3 Procedure

The experimental procedure and instructions given to participants were tightly constrained, to ensure that explicit learning strategies and receptive language difficulties were minimised. The exact wording of the instruction is included in supplementary materials.

Participants were exposed to the CVC language for five minutes. During this time they were simply instructed to listen to the language and to pay attention to the order of the words (the exact script for the instructions is available as supplementary material). They were then tested by being asked to decide whether 48 individual sequences (Table 4-2) were correct (i.e. consistent with the artificial grammar) or incorrect (i.e. violated
the artificial grammar in some way). Participants were able to express their decision either by pressing a button on a keyboard or custom made response box, or by pointing to yes or no on a piece of paper; whichever they found easier. At the end of a run, general overall feedback was provided with smiley to sad faces (Wong and Baker, 1988) according to overall percentage correct, along with the performance descriptor ‘Great!’ (>60%), ‘Well’ (55-60%), ‘OK’ (45-55%), or ‘Badly’ (<45%). This exposure-test cycle was then repeated in an identical fashion for the tone language. Again, they were explicitly instructed that the important thing was the order of the sounds.

After a short break, participants were then re-exposed to the CVC language for three minutes, before being re-tested. Sequences for both exposure and testing were presented in a different fixed pseudo-random order on each repetition. At the end of this run, feedback was provided relative to the previous CVC run with faces paired with the descriptors ‘Much Better’ (>110% of previous score), ‘Better’ (105-110%), ‘Same’ (90-105%), or ‘Worse’ (<90%). This procedure was then repeated for the tone language.

After a longer break, during which tea and biscuits were provided, participants completed a personal details questionnaire, which included questions about musical training and handedness (all patients were right handed). Patients then undertook the short form of the Boston Diagnostic Aphasia Examination (Kaplan, 1983) and the first half of the sentence comprehension section of the Verbs and Sentences Test to assess receptive grammar (Bastiaanse et al., 2003); controls completed an ACE-R and were tested on matrix reasoning (Raven, 1960); similarly demanding tasks of similar duration. Participants then completed another three minute exposure and test session on the CVC and tone languages, for a total of three testing sessions for each language.

Finally, each participant was re-exposed to the CVC language for three minutes, but the testing session that followed was replaced with an ‘oddball’ task. Participants were told that in this final test the ‘incorrect’ sequences were wrong in a different way, but were not explicitly instructed that they were looking for novel tokens. Where an ordering violation would first have occurred in an ‘incorrect’ sequence, the CVC token was replaced by a novel, previously unheard, oddball element (‘fen’, ‘muz’, ‘rol’, ‘dut’, ‘boz’ or ‘cav’). In this way, I was able to assess whether differential performance on the
three grammatical rule types was related to other undesired effects such as stimulus ordering.

All individuals undertook all study procedures on a single day, to ensure that differential patterns of performance consolidation during sleep did not confound my findings. The study procedures took up to four hours, including breaks.

Patients also undertook the short form of the Boston Diagnostic Aphasia Examination (Kaplan, 1983) and the first half of the sentence comprehension section of the Verbs and Sentences Test to assess receptive grammar; (Bastiaanse et al., 2003).

4.3.4 Stroke lesion mapping

The lesioned area of each brain was manually defined on every slice of each patient's 3T T1-weighted MRI scan in FSL, resulting in a 3D lesion mask. The resulting image was then registered to the standard MNI152 brain using FLIRT (FMRIB's Linear Image Registration Tool (Jenkinson and Smith, 2001); affine transformation model, 12 degrees of freedom). This registration matrix was used to register the patient's lesion mask to the standard space, from which a standardised lesion volume was computed in Matlab. Regions of interest in the left hemisphere (frontal inferior trigone, frontal inferior operculum, rolandic operculum, putamen and caudate) were identified using the aal atlas in SPM12, and the percentage of each sub-region that was lesioned was extracted for analysis (Table 4-3).
Table 4-3 Single subject lesion percentages by region for the stroke group.

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>Broca WM</th>
<th>Broca GM</th>
<th>FOP WM</th>
<th>FOP GM</th>
<th>Frontal Inferior Trigone</th>
<th>Frontal Inferior Operculum</th>
<th>Rolando Operculum</th>
<th>Putamen</th>
<th>Caudate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>99</td>
<td>98</td>
<td>99</td>
<td>97</td>
<td>98</td>
<td>98</td>
<td>92</td>
<td>74</td>
<td>28</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>33</td>
<td>30</td>
<td>19</td>
<td>33</td>
<td>22</td>
<td>46</td>
<td>77</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>45</td>
<td>0</td>
<td>34</td>
<td>83</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>77</td>
<td>86</td>
<td>77</td>
<td>89</td>
<td>84</td>
<td>86</td>
<td>89</td>
<td>70</td>
<td>42</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>96</td>
<td>94</td>
<td>61</td>
<td>97</td>
<td>70</td>
<td>71</td>
<td>12</td>
<td>59</td>
</tr>
<tr>
<td>6</td>
<td>89</td>
<td>91</td>
<td>85</td>
<td>89</td>
<td>91</td>
<td>88</td>
<td>100</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>21</td>
<td>3</td>
<td>0</td>
<td>8</td>
<td>26</td>
<td>51</td>
<td>36</td>
</tr>
<tr>
<td>8</td>
<td>54</td>
<td>32</td>
<td>97</td>
<td>73</td>
<td>38</td>
<td>80</td>
<td>99</td>
<td>76</td>
<td>10</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>1</td>
<td>35</td>
<td>8</td>
<td>5</td>
<td>16</td>
<td>0</td>
<td>59</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>79</td>
<td>59</td>
<td>92</td>
<td>62</td>
<td>65</td>
<td>70</td>
<td>48</td>
<td>48</td>
<td>46</td>
</tr>
<tr>
<td>11</td>
<td>32</td>
<td>5</td>
<td>30</td>
<td>16</td>
<td>20</td>
<td>13</td>
<td>21</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>99</td>
<td>98</td>
<td>99</td>
<td>97</td>
<td>98</td>
<td>98</td>
<td>92</td>
<td>74</td>
<td>28</td>
</tr>
</tbody>
</table>

4.3.5 nfvPPA atrophy mapping

Nine of the patients in the nfvPPA group underwent a 3T volumetric T1 MRI scan. From a database of healthy control scans from the same scanner, an age-matched normative sample of 36 individuals was selected by finding the four nearest-neighbours to each patient in terms of age, excluding duplication (mean age of these controls was 73 years). After segmentation, the nine nfvPPA scans and nine nearest-neighbour controls were used to create a DARTEL template. This was then applied to the remaining 27 controls. Resultant images were normalised to MNI space in SPM12 with an 8mm smoothing kernel, and separate statistical comparisons were performed for grey and white matter, with total intracranial volume and age as covariates.

As for the stroke lesion mapping, regions of interest in the nfvPPA patients were then identified using the aal atlas, normalised to MNI space. As distinct from the lesion analysis in the stroke cases, these regions were defined bilaterally. The grey matter volume in each region was then extracted by applying these regions to each individual’s modulated, warped grey matter segmentation and correcting for total intracranial volume.

4.3.6 Analysis

Performance metrics for analysis were based on signal detection theory (Stanislaw and Todorov, 1999). Standard signal detection measures of discriminability and bias rely
on the underlying trial difficulty within a run being constant. In my artificial grammar, it was expected that violations of the linear relationship (complexity level 1) would be easier to detect than configurational violations (complexity level 2), which in turn might be easier to detect than violations of the hierarchical structure (complexity level 3) (Figure 4-5). To accommodate this, I separately calculated d’ and the non-parametric equivalent A’, based on a comparison of performance on each of the three types of violation sequence to all of the consistent sequences. A single value for bias measures was also calculated based on the combined distribution of violation sequences (Figure 4-5). Hit and false alarm rates of 1 were replaced with \((n-0.5)/n\) (where \(n\) is the number of trials), and those of 0 with \(0.5/n\) (Macmillan and Kaplan, 1985, Stanislaw and Todorov, 1999). The non-parametric analogues of these signal detection metrics, A’ for discriminability and \(\beta''\) for bias, were used for the primary analysis.
Figure 4-5 Distribution of decision variables. Cartoon illustrating the distribution of parametric decision variables for an hypothetical experiment with easy, medium and hard to detect rule violations. $d'$ represents the single subject discriminability of each rule violation, while $c$ and $\beta$ represent different measures of bias (i.e. in this case the tendency to say that a sequence is grammatical if there is no evidence to the contrary). Each rule violation has its own $d'$ measure, reflecting its respective discrimination difficulty, while the bias measures apply to the experimental context as a whole.
All statistical analyses were performed in Matlab R2015b with the Statistics and Machine Learning Toolbox unless otherwise specified. Differences from chance performance in both discriminability measures (A’ and d’) and measures of bias (ln(β), c and β”) were assessed for each group and condition separately using one-sample Wilcoxon signed rank tests (the non-parametric equivalent of the one sample t-test).

The effects of group and rule complexity on discriminability were assessed with three separate repeated measures ANOVA tests (one for each test type; CVC, tones and oddball), with the factor ‘participant number’ nested within ‘group’. This parametric statistical test was employed because there is no appropriate non-parametric test for repeated measures designs of the kind employed here; the Friedman test does not allow multiple groups to be compared. Significant results were explored with post-hoc comparisons of population marginal means.

The degree of learning across exposure-test pairs was assessed by fitting a general linear model in Minitab 17 for the CVC and tone languages. The response variable was discriminability (A’) and the factors were ‘participant number’ (nested within ‘group’), ‘rule type’, and ‘run number’. Significant results were explored with post-hoc Tukey’s range tests.

To assess whether the same rules were learned by participants between task types (CVC vs tones vs oddball), and by extension whether learning of the same artificial grammar was transferrable across token types (for CVC vs tones), Spearman correlation matrices were constructed based on performance patterns by sequence for each group. From these, hierarchical cluster analysis was performed to construct dendrograms representing the similarity of performance pattern across test and group, using a ‘farthest neighbour’ linkage method with a data-driven inconsistency coefficient (The Mathworks Inc., 2015).

Exploratory regression analyses were performed in Minitab 17. As I measured a large number of variables for each subject, I performed stepwise regression to determine those variables that best predicted artificial grammar learning. This is an automated process to identify a useful subset of predictors by sequentially adding and removing predictors until an optimal model is obtained. Software default alpha-to-enter and alpha-to-remove values of 0.15 were used, with confirmatory analyses at 0.1/0.15 and
0.1/0.1 yielding identical results. Three separate sets of stepwise regressions were performed; one to explore possible correlations between rule discriminability and neuropsychological and language measures, a second to explore correlations between rule discriminability and stroke lesion site, and a third to explore correlations between rule discriminability and grey matter volume in nfvPPA. The potential continuous predictors included in the first model set were Age, Raven’s Progressive Matrix score, years of musical training (which I hypothesised might impact tone language difficulty), sentence comprehension (from the Verbs and Sentences Test), and overall aphasia severity (Aphasia Severity Rating Score). For the second set, the potential continuous predictors were age, the proportion of each region of interest lesioned, total lesion volume, and the number of years since stroke. For the third, potential predictors were age, grey matter volume summed across regions of interest in each hemisphere, and corrected whole brain grey matter volume.

4.4 Results

4.4.1 Discriminibility of sequencing rules

Raw performance for each individual sequence is shown in Figure 4-6 A, and the results of the signal detection theory analysis are illustrated in Figure 4-6 B. The hierarchical cluster analysis is illustrated in Figure 4-8. Performance did not differ between sequences heard during exposure and test phases (Figure 4-6 A sequences 1-3) and those that were novel during the test phase (Figure 4-6 A sequences 4-6), so these were collapsed. Results are presented for the non-parametric discrimination measure A’, but the same pattern of findings was present for the parametric equivalent d’ (see appendix Table 4-5, Table 4-6).
Figure 4-6 Group performance on sequence identification. Dashed lines represent chance performance. Error bars represent group-wise standard error of the mean. Controls are in blue, nfvPPA in red and stroke in orange. A) Proportion of correct responses for each testing sequence by group and task. Sequences 1-3 were consistent with the grammar and familiar from the exposure phase, while 4-6 were consistent and novel. Sequences 7-12 contained violations of the types indicated (see table 2). B) Discriminability of each rule type by group for each language type. C) Overall discriminability by group and run number for the CVC language (improving performance by run represents learning over time). D) Overall discriminability by group and run number for the tone language.
<table>
<thead>
<tr>
<th>A</th>
<th>Rule complexity</th>
<th>Group</th>
<th>Group x complexity</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC</td>
<td>F(2,60) = 14.8</td>
<td>F(2,60) = 4.1</td>
<td>F(4,60) = 0.7</td>
<td>F(30,60) = 1.4</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.0001</td>
<td>p=0.0264</td>
<td>p=0.5933</td>
<td>p=0.1246</td>
</tr>
<tr>
<td>Tones</td>
<td>F(2,60) = 9.1</td>
<td>F(2,60) = 7.1</td>
<td>F(4,60) = 0.5</td>
<td>F(30,60) = 0.3</td>
</tr>
<tr>
<td></td>
<td>p=0.0004</td>
<td>p=0.0029</td>
<td>p=0.7421</td>
<td>p=0.9995</td>
</tr>
<tr>
<td>Oddball</td>
<td>F(2,60) = 0.5</td>
<td>F(2,60) = 2.6</td>
<td>F(4,60) = 0.2</td>
<td>F(30,60) = 11.9</td>
</tr>
<tr>
<td></td>
<td>p=0.6018</td>
<td>p=0.0916</td>
<td>p=0.9575</td>
<td>p&lt;0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>Bias metric</th>
<th>Group</th>
<th>Group x metric</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC</td>
<td>F(2,60) = 6.7</td>
<td>F(2,60) = 2.3</td>
<td>F(4,60) = 0.5</td>
<td>F(30,60) = 4.0</td>
</tr>
<tr>
<td></td>
<td>p=0.0024</td>
<td>p=0.1150</td>
<td>p=0.7721</td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td>Tones</td>
<td>F(2,60) = 19.1</td>
<td>F(2,60) = 0.1</td>
<td>F(4,60) = 1.0</td>
<td>F(30,60) = 1.7</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.0001</td>
<td>p=0.9445</td>
<td>p=0.4184</td>
<td>p=0.0371</td>
</tr>
<tr>
<td>Oddball</td>
<td>F(2,60) = 3.6</td>
<td>F(2,60) = 0.2</td>
<td>F(4,60) = 0.6</td>
<td>F(30,60) = 8.2</td>
</tr>
<tr>
<td></td>
<td>p=0.033</td>
<td>p=0.8344</td>
<td>p=0.6516</td>
<td>p&lt;0.0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C</th>
<th>Rule</th>
<th>Run</th>
<th>Group</th>
<th>Run x Complexity</th>
<th>Run x Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC</td>
<td>F(2,890) = 141.2</td>
<td>F(2,890) = 28.5</td>
<td>F(30,890) = 14.4</td>
<td>F(4,890) = 0.9</td>
<td>F(4,890) = 0.7</td>
</tr>
<tr>
<td></td>
<td>p&lt;0.001</td>
<td>p&lt;0.001</td>
<td>p=0.029</td>
<td>p=0.626</td>
<td>p=0.479</td>
</tr>
<tr>
<td>Tones</td>
<td>F(2,890) = 73.8</td>
<td>F(2,890) = 2.9</td>
<td>F(30,890) = 4.3</td>
<td>F(4,890) = 0.7</td>
<td>F(4,890) = 3.3</td>
</tr>
<tr>
<td></td>
<td>p=0.001</td>
<td>p=0.055</td>
<td>p=0.026</td>
<td>p=0.572</td>
<td>p=0.010</td>
</tr>
</tbody>
</table>

Table 4-4 Statistical tests corresponding to Figure 4-6 and Figure 4-7.
A and B: Repeated measures ANOVAs of group against rule for the non-parametric discriminability measure A’ (panel A; corresponding to Figure 4-6 B) and bias (panel B; corresponding to Figure 4-7), with participant number as a nested factor within group. C: The general linear model assessing learning across runs for the non-parametric discriminability measure A’ (corresponding to Figure 4-6 C (CVC) and Figure 4-6 D (Tones)).

For the CVC language, discriminability as measured by A’ showed significant main effects of rule complexity and group, but no group x rule interaction or consistent inter-individual differences (Figure 4-6 B, Table 4-4 A). Post-hoc comparison of marginal means indicated that the group difference is driven by the control participants performing significantly better than participants in the stroke group (p=0.0058). Participants with nfvPPA performed at an intermediate level, and were not statistically different from either controls (p=0.14) or stroke (p=0.50). Further, all groups performed better at detecting violations of linear grammatical rules than configurational (p=0.0001) or hierarchical (p=0.0001), which in turn did not differ (p=0.99).

The tone language discriminability measured by A’ showed significant main effects of rule complexity and group, but no group x rule interaction or consistent inter-individual differences (Figure 4-6 B, Table 4-4 A). For this language, post-hoc
comparison of marginal means indicated that the group difference is driven by the control participants performing significantly better than participants in the nfvPPA group (p=0.048). Participants with stroke performed at an intermediate level, and were not statistically different from either controls (p=0.31) or nfvPPA (p=0.57). Rather than showing lower performance with increasingly complex rule violations, all groups performed significantly more poorly at detecting violations of the configurational rules than both the linear (p=0.0009) and hierarchical (p=0.0001), which in turn did not differ (p=0.73). Possible reasons for this unexpected pattern are discussed below.

All groups performed well at discriminating the oddball stimuli. For all groups, mean and median discriminability was better than for the CVC or Tone languages. Crucially, there was no effect of rule type for the oddball language (Figure 4-6 B, Table 4-4 A). This is because violations were no longer based on detection of grammatical rules, but on the detection of novel CVC tokens in various positions within the sequence. This observation reassures us that the effect of rule complexity in the CVC language cannot be explained by the position of the violation within a sequence. There was no significant group difference in performance or group x rule interaction, but there was a strongly significant effect of participant, indicating that individuals within each group differed in their ability to detect the novel non-word tokens. Post-hoc comparison of marginal means was not performed, as there was no rationale from the ANOVA to proceed to this.

4.4.2 Bias towards permissibility

From Figure 4-6 A it can be seen that for some sequences, particularly those that violate the more complex rules, participants in all groups performed at a level below chance (50% correct). This would not be expected if participants were simply guessing for these sequences, but would be expected if the grammatical violation was sufficiently subtle that participants made an active decision that the sequence is not inconsistent with the grammar. In other words, participants would display bias towards stating that a sequence was consistent with the artificial grammar unless they had evidence otherwise, but it was not known whether this tendency would differ between groups (Haendiges et al., 1996, Dickey et al., 2008). A bias towards yes was demonstrated,
but there was no group difference in bias or group x metric interaction (Figure 4-7, Table 4-4 B).

Figure 4-7 Bias metrics by group for each language type
Values less than zero indicate a bias towards classifying sequences as correct.

4.4.3 Implicit learning through repeated exposure-test cycles

Participants, including patients, improved with practice. Results of the general linear model analysis including run number are illustrated in Figure 4-6 C, Figure 4-6 D, and Table 4-4 C. For the CVC language, there were main effects of rule type, run number and group, but no interactions between run number and either rule complexity or group. This suggests, (a) that all groups were learning across repeated exposure-test cycles, (b) that the amount of learning over time was not different between the three rule types, and (c) that the group difference in overall discriminability described above was driven by differences in initial grammatical learning, not by a reduced ability to refine the internal grammatical model through feedback and repeated exposure. Tukey tests confirmed that participants improved on each set of exposure and testing (p<0.05 in all cases). Mean A’ across groups for run 1 was 0.55, for run 2 was 0.60, and for run 3 was 0.64 (chance performance is 0.5). Control performance was significantly superior to nfvPPA performance, which was in turn significantly superior to performance in the stroke group.

For the tone language, there were main effects of rule type and group, but only a trend towards a main effect of run number. There was an interaction between run number and group, but not run number and rule complexity. Tukey tests demonstrated that control participants performed significantly better than those with stroke, who in turn performed significantly better than those with nfvPPA. Performance on run 3 was significantly better than on run 1, while run 2 did not differ from either run 1 or 3.
A) Correlation Matrix

B) Spearman Linkage Dendrogram

Figure 4-8 Correlation and clustering analysis. 
A) Spearman rank-order correlation matrices. B) Linkage based on hierarchical cluster analysis of Spearman correlations. Three clusters emerge with a linkage distance cut-off of 0.5, and are indicated in colour groupings (blue, green and red).

4.4.4 Separation of rule learning between linguistic and non-linguistic sequences

Spearman rank-order correlograms are shown in Figure 4-8 A, and the resultant hierarchical cluster analysis is visualised in Figure 4-8 B. Rule acquisition patterns are highly correlated between groups within each task, but not between tasks within each group. This is confirmed by the cluster analysis. Performance on CVC and Tone languages represent strongly distinct clusters, while Oddball performance profiles are
less distinct (as expected if sequence ordering is unimportant for novel token detection). The difference in profiles between CVC and Tone languages across group was further assessed by a two-sample t-test with unequal variance based on the similarities shown in Figure 4-8 A. In sample 1 were the six within-language similarities (excluding the diagonal) and in sample 2 were the 9 between-language similarities. This confirmed a highly significant group difference in language similarity; t(12.5)=5.2, p=0.0002. Identical pair-wise tests between groups across language (blinded to overall ability by non-parametric Spearman rank-order correlation) confirmed that performance profiles did not differ between groups (control vs nfvPPA t(12.3)=−0.12, p=0.90; control vs stroke t(10.2)=−0.76, p=0.46; nfvPPA vs stroke t(12.2)=−0.01, p=0.99).

4.4.5 Relationship to diagnostic category and disease severity

Stepwise regression analysis between overall measures of discriminability and the behavioural measures listed in the methods yielded no statistically significant predictors for CVC or oddball language performance. If performance on CVC linear rules (where performance was highest) is considered in isolation, overall aphasia severity (α=0.002) was the only significant predictor. The only significant predictor for tone language performance was diagnosis (α=0.035), confirming that patients with nfvPPA performed more poorly than those with stroke, independent of aphasia severity.

4.4.6 Relationship to structural anatomy

Lesion volume and site affected performance in the stroke group. Stepwise regression analysis between overall measures of discriminability for each language and the lesion metrics listed in methods yielded a model for the CVC language including only left putamen (α=0.06); in other words, the ability to detect sequencing violations decreased with more severe putaminal lesions. For the oddball language, total lesion volume (α=0.016) and involvement of the left ventral frontal operculum (α=0.076) were included in the model (overall p=0.023), but in opposite directions: participants with larger lesions were less able to detect oddball CVCs, but this deficit was ameliorated if
their lesion had a more anterior distribution. The model yielded no statistically significant predictors for the tones language.

Grey matter volume affected performance in the nfvPPA group. The best model for nfvPPA performance on the CVC language included age ($\alpha=0.002$) and total grey matter volume in the left frontal lobe regions of interest ($\alpha=0.016$), but not in their right sided equivalents or total grey matter volume. These acted such that performance improved with higher left frontal grey matter volume, and also with age. While it might initially seem counter intuitive that older patients performed better, this is likely to reflect the natural loss of grey matter with age. The model is therefore improved by accounting for the fact that any given value of grey matter volume is relatively more atrophic in a younger individual. There were no statistically significant predictors with grey matter volume in the nfvPPA group for performance on the tone or oddball languages.

4.5 Discussion

This study successfully used a mixed-complexity artificial grammar learning task with speech sounds and tone stimuli to test aphasic patients with two different aetiologies. The principal observations were that: 1) both healthy individuals and patients with aphasia apply strongly contrasting strategies to assess structured sequences depending on whether the sequences consist of linguistic or non-linguistic auditory tokens; 2) patients with vascular aphasia and nfvPPA show similar patterns of auditory sequence processing impairment compared to controls; 3) aphasic patients are not disproportionately impaired on more complex auditory sequencing tasks, instead displaying a general impairment in processing structured auditory input; 4) patients with aphasia are capable of implicit learning of this kind through repeated exposure/test cycles. I discuss these results in turn in the following sections.

4.5.1 Rule acquisition differs when structurally identical sequences are comprised of linguistic or non-linguistic stimuli.

In all groups, performance profiles on the CVC language followed the expected pattern of linear relationships being more discriminable than configurational or hierarchical
structures. By contrast, participants' judgments about the tone language did not seem to be based on the abstraction of the intended grammatical rules (Figure 4-6 A, panel 2). This impression was confirmed by correlation and cluster analyses (Figure 4-8). Hierarchical clustering based on non-parametric correlations of single subject performance profiles was clearly able to recover the language learned, but not the group structure. This demonstrates that all groups acquired the same set of rules when making decisions about the CVC language; all that differed between groups was their overall performance. Further, a completely different set of rules were acquired for the tone language, but again this learning profile was almost identical across groups. Therefore, rule acquisition was not transferred between the two languages, and the separation of approach to linguistic and non-linguistic structured sequences was strongly maintained despite agrammatic aphasia of either type. This is despite the two languages having an identical structure, and being presented and tested in the same order. This finding cannot be trivially explained by a lower level deficit such as the tone language simply being more difficult, more affected by a reduced fidelity of auditory processing or subject to a higher ‘lapse rate’, which would affect discriminability but could not produce the complete dissociation of response patterns shown here across all groups.

Despite extensive exposure, all groups performed poorly at classifying tone sequence number 1 as consistent with the artificial grammar, and indeed seemed to actively reject it, with control performance for this sequence well below chance. This sequence was comprised entirely of tone sweeps, embedded within a recursive structure. The participants also correctly classified sequence 12, which has similar properties, as inconsistent with the artificial grammar. In this case, good performance on this tone sequence does not reflect an ability to extract the hierarchical rule, but rather a consistent tendency to reject the embedded pattern of tone sweep sequences. Overall, therefore, it does not seem that the tone sequences were assessed for the specific violations of the grammatical rules inherent in the artificial language. Instead, they were judged on the overall ‘feel’ of the sequence in a manner very different to the CVC language but entirely consistent across groups.

I therefore infer that, while both patient groups are impaired in their ability to learn and discriminate sequencing rules for both CVC and tone languages, they maintain the
same separation of processing of these languages seen in control participants. Taken together, these results imply that domain specific processes exist for linguistic and non-linguistic structured sequence learning, which are preserved even in the presence of acquired grammatical deficits. These processes might therefore engage different brain networks (Geranmayeh et al., 2014b). It is possible that this separation is instantiated by an assessment of phonological ‘well-formedness’ in auditory temporal regions (Obrig et al., 2016).

4.5.2 Artificial grammar learning in aphasia is similar across aetiologies.

A second key observation was that patients with nfvPPA and stroke showed similar patterns of performance for the CVC language. All groups were able to correctly classify the grammatical testing sequences as consistent with the exemplary sequences heard during exposure, and this ability fully generalised from the exposure set to the novel sequences not heard during exposure (Table 4-2; Figure 4-6 A). While there was a group difference, with performance in the stroke group being poorer on the CVC task, this effect disappeared when overall aphasia severity was accounted for. This novel result in agrammatic patients with primary progressive aphasia provides the evidence to suggest that existing findings on stroke patients should prove applicable to the progressive aphasias.

The lack of an effect of rule type for the oddball task confirms that the pattern of performance for the CVC language cannot be explained by stimulus-level differences such as the position of the violation within the sequence. The lack of a statistical group difference in the ability to detect oddball CVCs, a task performed at the end of the testing session, suggests that the patients’ impairments in this study are not solely due to generic difficulties with performing psychophysical sequence processing tasks, latent yes/no confusion, difficulties with basic auditory processing or differential effects of fatigue between groups.

The only consistent group difference not accounted for by severity was that patients with nfvPPA performed more poorly than those with stroke on the tone based language. There are a number of possible reasons for this. It might be a consequence of the tone language being more affected by the basic auditory processing deficits
previously demonstrated in nfvPPA (Goll et al., 2010, Grube et al., 2016). Alternatively, it might reflect involvement of the right IFC, which was spared in the stroke group (Figure 4-1 B), and is posited to have a role in prosodic and tonality based judgments. Nonetheless, patients with nfvPPA maintained the same pattern of learning as the other groups (Figure 4-8), demonstrating that this deficit is a specific difficulty with processing tonal input rather than a breakdown of the separation of phonological vs tonal structured sequence processing.

4.5.3 Agrammatic aphasic patients are similarly impaired for both complex and simple sequencing operations.

As expected, patients did not perform as well as controls, but the magnitude of this performance deficit did not differ by rule complexity, counter to my initial hypothesis. The results demonstrate that aphasic patients showed a global deficit in sequence processing, and were not selectively impaired on complex sequences. Clinically, patients in both groups make errors in the parsing of more complex syntax, and tend to stick to active, subject-relative structures in their expressive language (Grossman and Moore, 2005). I suggest that this does not reflect a specific deficit in the processing of more complex linguistic structures, but rather that these constructions are simply more difficult and therefore more vulnerable to a global deficit. As well as explaining a clinical symptom, this conclusion is consistent with the functional imaging finding that, in nfvPPA, left inferior frontal cortex activity lacks the normal relationship with syntactic complexity (Wilson et al., 2010); it suggests that the efficiency of IFC is so degraded that even the least complex grammatical structures require maximal neural recruitment. This is analogous to the finding that older adults are no longer able to selectively modulate anterior cingulate cortex in response to increasingly difficult listening environments as they have already fully engaged this region in easy listening conditions (Erb and Obleser, 2013).

It is possible that patients attempt to compensate for this deficit by engaging a wider syntactic processing network involving temporo-parietal regions (Schofield et al., 2012b, Blank et al., 2016), but that this is insufficient to compensate for lost language function (Wilson et al., 2016). In stroke, where the IFC is lost entirely, the presence of residual ability could be due to complete reliance on this wider network (Thompson et
al., 2010), or the involvement of contralateral IFC. Future functional imaging studies of implicit grammar learning will inform this debate. In any case, my findings of a general impact on sequence processing suggest that the grammatical deficits observed in aphasia, and the recovery from these, might reflect higher level, domain general processes (Caramazza and Zurif, 1976, Dominey et al., 2003, Patel et al., 2008, Brownsett et al., 2014, Geranmayeh et al., 2014a, Geranmayeh et al., 2017).

The lack of group by rule complexity interactions is unlikely to be explained by limitations in sample size, as in no case was there even a trend in this direction (all interactions with complexity in Table 4-4 have p-values > 0.5). Therefore, these results appear to represent a genuine null effect, rather than sub-threshold effects that might become significant with a greater sample size. Moreover, the lack of interaction is also consistent in the tone and CVC experiments, although the response patterns between the two are vastly different. Nor can the results be trivially explained by floor effects in processing the more complex relationships for the following reasons: 1) the nfvPPA group show strikingly parallel behaviour in relation to the control group, consistent with a proportional impairment even on the linear sequencing operation (Figure 4-6 B, panel 1); 2) the stroke patients may well have reached a floor in performance on the complex sequences in the nonsense word task, but even excluding this group did not cause the group by complexity interaction to approach significance; 3) all groups improved over the three testing runs, and performance improvement was parallel in relation to sequencing complexity (i.e. there was a main effect of run but no run-by-complexity interaction, Figure 4-6 C); and 4) the tone language showed a very different pattern of results to that for the nonsense words yet, again, there was no evidence of a group-by-complexity interaction.

4.5.4 Patients with aphasia show improved performance over repeated cycles.

All three groups demonstrated the same amount of learning across repeated exposure/test cycles of the CVC language; the only difference was in their initial levels of performance (Figure 4-6 C; Table 4-4 C). This suggests that patients with aphasia were able to update their internal model of the artificial grammar based on feedback
and implicit comparison with short periods of exposure. It also suggests that patients did not suffer greater effects of fatigue than controls. By contrast, learning did not occur to the same degree for the tone language (Figure 4-6 D; Table 4-4 C). It is therefore clear that learning was not transferrable between the CVC and tone languages, despite them sharing the same underlying artificial grammar structure.

Together, these findings provide a theoretical basis upon which an exposure-based speech therapy for grammar could be built with the aim of improving subjective difficulty with speech comprehension (Figure 4-2). The results imply that there is potential for improvement from an intensive paradigm based on repeated exposure-test cycles. This could in principle be made home-deliverable and patient-led, an approach that has demonstrable efficacy for improving speech production in similar patient groups (Varley et al., 2016); the tasks employed in this study were automated and computer based. The finding that learning did not generalise across modalities implies that such a therapy would need to use linguistic material, as it would be unlikely to be so well learnt with non-linguistic material or to transfer across domains. In contrast, the finding that patients were able to generalise perfectly from sequences heard during exposure (Figure 4-6 A, sequences 1-3), to those that were novel during the test phase (Figure 4-6 A, sequences 4-6), to the extent that performance did not differ, suggests that such a therapy might not need to be comprehensive with regards to specific sentence structures. Instead, it is envisaged that a graded programme could be designed, such that training focusses initially on those structures that are having most frequent impact on speech comprehension. Clearly my study does not provide evidence that such a therapy would be more efficacious than existing methods for addressing asyntactic deficits after stroke, nor do I have any evidence of how well these strategies would work within an already-learned but now-impaired natural language. Indeed, a recent small study of nine patients who had chronic agrammatic aphasia secondary to stroke suggests that implicit learning alone (on a visuo-motor serial reaction time task) does not necessarily translate into improved real-world performance (Schuchard et al., 2017). Larger scale trials and assessment with other tasks are clearly required in more acute disease cohorts. As my method relies on exposure-based implicit learning rather than explicit instruction, future therapeutic trials could follow recent trends for patient-led practice in the home environment, increasingly with support from internet-based...
resources (Rogalski et al., 2016). This has the potential for providing the benefits of an intensive approach (Bhogal et al., 2003, Brady et al., 2016, Breitenstein et al., 2017) without the resource constraints that limit the frequency and therefore efficacy of traditional irregular, face-to-face instruction (Sarno et al., 1970, Lincoln et al., 1984).

4.5.5 Brain behaviour relationships

My use of stepwise regression, to assess associations between brain structure and function, should be seen as exploratory, since the study was powered only to detect strong effect sizes. Nonetheless, significant associations between brain structure and behavioural performance within the patient groups were observed. The nfvPPA group demonstrated significant grey and white matter loss in frontal regions bilaterally. In keeping with previous studies of expressive grammar (Rogalski et al., 2011), the discriminability of grammatical structure in the CVC language correlated with age-corrected loss of volume only in left frontal regions (but not similar right-sided regions or total corrected grey matter volume). No such association was found for the tone language or the oddball task.

In the stroke group there was catastrophic loss of fronto-temporal regions in the dominant hemisphere, but complete contralateral preservation. For the CVC language, the model included only the putamen. The putamen is known from the functional imaging literature to be important in implicit sequence detection and learning in healthy individuals (Grafton et al., 1995, Rauch et al., 1995, Rauch et al., 1997, Grahn and Rowe, 2009). No such relationships were significant for the tone task. Performance on the oddball sound detection task was predicted by overall lesion volume and a weaker, opposite, effect of the extent of involvement of the most frontal region analysed. This suggests that more anterior lesion locations were less deleterious to CVC Oddball detection than those located closer to auditory temporal cortex.

Similar performance in a relatively bilateral disease (nfvPPA) and one so clearly unilateral (stroke) immediately poses the question of which components of a broader language network are recruited to underpin the demonstrated learning over repeated exposure-test cycles (Crinion and Price, 2005). Thus, there is clear scope for a
functional imaging study to be conducted in these patient groups, the results of which could complement the development of grammar-based speech therapy.

4.5.6 Study limitations

The main limitation of this study is the relatively small number of patients in each group, which renders the correlation analyses exploratory, and more subtle voxel-based lesion-symptom mapping analyses impossible. Future work in larger and more heterogeneous cohorts would complement functional neuroimaging in exploring the brain basis of implicit sequence learning.

The stroke group all had left-hemisphere lesions, while the nfvPPA group had relatively bilateral disease that involved a more distributed language network. I have demonstrated a preserved separation of rule learning between the CVC and tone languages in both disease groups, confirming the domain specificity of structured sequence learning. However, overall performance on the tone language was poorer in nfvPPA, and the reasons for this are not clear. The lack of a cohort of stroke patients with right-hemisphere lesions, auditory psychophysical examination of all groups, or detailed neuropsychological testing including an assessment of working memory, means that discussion of the potential brain basis for this difference is speculation that must be confirmed or refuted in future work.

4.6 Conclusion

In this chapter I reconcile a controversy in the literature regarding the effects of structural complexity on receptive grammar in the frontal aphasias. I demonstrate that, while the patients found complex, non-adjacent structuring relationships more difficult to acquire, this did not represent a disproportionate impairment; aphasia resulted in a similar performance penalty for adjacent relationships. I also provide insights into the language-specificity of artificial grammar learning by demonstrating that humans learn otherwise identical linguistic and non-linguistic structured sequences entirely separately, even if the neural architecture underlying this learning is disrupted. My direct comparison of two patient groups suggests that previous findings regarding implicit sequence learning in stroke aphasia are likely to prove transferrable to nfvPPA.
Finally, the ability of both patient groups to learn an artificial grammar as demonstrated here provides a rationale and approach for future trials of implicit, exposure-based approaches to rehabilitation of agrammatism in non-fluent aphasia.
4.7 Appendix: Supplementary tables

<table>
<thead>
<tr>
<th>A</th>
<th>Rule complexity</th>
<th>Group</th>
<th>Group x Rule</th>
<th>Participant</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC</td>
<td>&lt;0.0001</td>
<td>0.0311</td>
<td>0.6445</td>
<td>0.0006</td>
</tr>
<tr>
<td>Tones</td>
<td>0.0006</td>
<td>0.0027</td>
<td>0.6069</td>
<td>0.9994</td>
</tr>
<tr>
<td>Oddball</td>
<td>0.1383</td>
<td>0.0611</td>
<td>0.7245</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Table 4-5 Supplementary analysis of d’ discriminability
P-values for a repeated measures ANOVA of group against rule for the parametric discriminability measure d’, with participant number as a nested factor within group.

<table>
<thead>
<tr>
<th>A</th>
<th>Rule</th>
<th>Run</th>
<th>Group</th>
<th>Rule x Run</th>
<th>Run x Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVC</td>
<td>&lt;0.00</td>
<td>&lt;0.00</td>
<td>0.020</td>
<td>1</td>
<td>0.054</td>
</tr>
<tr>
<td>Tones</td>
<td>0.070</td>
<td>&lt;0.00</td>
<td>0.010</td>
<td>0.333</td>
<td>0.019</td>
</tr>
</tbody>
</table>

Table 4-6 Supplementary analysis of d’ learning
P-values for the general linear model assessing learning across runs for the parametric discriminability measure d’.
4.8 Chapter Summary:

Patients with non-fluent aphasias display impairments of expressive and receptive grammar. This has been attributed to deficits in processing configurational and hierarchical sequencing relationships. This hypothesis had not been formally tested. It was also controversial whether impairments are specific to language, or reflect domain general deficits in processing structured auditory sequences.

Here I used an artificial grammar learning paradigm to compare the abilities of controls to participants with agrammatic aphasia of two different aetiologies: stroke and frontotemporal dementia.

Patients with non-fluent variant primary progressive aphasia (nfvPPA), non-fluent aphasia due to stroke, and controls implicitly learned a novel mixed-complexity artificial grammar designed to assess processing of increasingly complex sequencing relationships. I compared response profiles for otherwise identical sequences of speech tokens (nonsense words) and tone sweeps.

In all three groups the ability to detect grammatical violations varied with sequence complexity, with performance improving over time and better for adjacent than non-adjacent phonological relationships. Patients performed less well than controls overall, and this was related more strongly to aphasia severity than to aetiology. All groups improved with practice and performed well at a control task of detecting oddball nonwords. Crucially, group differences did not interact with sequence complexity, demonstrating that aphasic patients were not disproportionately impaired on complex structures. Hierarchical cluster analysis revealed that response patterns were very similar across all three groups, but very different in the nonsense word and tone tasks, despite identical artificial grammar structures.

Overall, I demonstrate that agrammatic aphasics of two different aetiologies are not disproportionately impaired on complex sequencing relationships, and that the learning of phonological and non-linguistic sequences occurs independently. The similarity of profiles of discriminatory abilities and rule learning across groups suggests that insights from previous studies of implicit sequence learning in vascular aphasia are likely to prove applicable in nfvPPA.
Chapter 5: Review and Synthesis

5.1 Network organisation and reorganisation in neurodegeneration

Neurons within the human brain are functionally and structurally connected at a number of scales. With neuroimaging it is possible to assess this connectivity at the mesoscopic scale, visualising the way in which groups of neurons interact with each other across brain regions. In chapter 2, I observed such connectivity in terms of shared temporal dynamics in fMRI signals (Hampson et al., 2002). In chapter 3, I went beyond this by taking direct measurements of neuronally-derived magnetic and electrical signals at high temporal resolution, with simultaneous magnetoencephalography and electroencephalography. This allowed me to test for the presence of consistent phase relationships between distant brain regions, manifesting as signal coherence (Nolte et al., 2004). Having thus demonstrated functional connectivity (statistical dependency), I assessed effective connectivity (causal interaction) by fitting a frequency-resolved multivariate autoregressive model to these same data, which then allowed me to directly evaluate the Granger-causal relationships between regions (Granger, 1969, Gow et al., 2008).

By combining measurements of functional connectivity with experimental psychology in chapter 3, I robustly demonstrated one way in which the connectivity of the healthy brain underpins perception and higher cognitive function by allowing the rapid and flexible integration of predictions with sensory input. By taking a cross-sectional approach to multimodal imaging data in chapter 2, I have also demonstrated the way in which connectivity contributes to the progression of neurodegeneration through trans-neuronal spread of tau pathology in Alzheimer’s disease and regional vulnerability factors in progressive supranuclear palsy, settling a key debate about reasons for differential network vulnerability in Alzheimer’s disease and fronto-temporal lobar degeneration (Zhou and Seeley, 2014). Therefore, while the methods employed in chapters 2 and 3 are different, they provide complementary insights into the interplay between brain network organisation, function and degeneration.
As well as exploring the nature, purpose and effects of healthy brain network organisation, I have precisely defined the consequences of regional neurodegenerative burden for brain networks at the local and global scale. In chapter 2, by employing a cross-sectional approach, I was able to reconcile previous literature that seemed to provide contradictory evidence about the consequences of Alzheimer’s disease pathology for functional connectivity. Although based on a relatively small number of individuals (n=17), my data were consistent with computational models that predict a dissociation between changes in functional connectivity in early and late Alzheimer’s disease (de Haan et al., 2012). In this framework, neurons initially compensate for synaptic loss by increasing their firing rate, manifesting as stronger functional connectivity (Maestú et al., 2015). As neurodegeneration progresses, this mechanism breaks down, and functional connectivity weakens (Jones et al., 2015). This state of affairs can account for group-average results demonstrating either weakened or strengthened connectivity depending on the stage of illness (and hence position on the ‘inverted U’ shaped curve) at which patients were assessed.

In PSP, I demonstrated diaschisis, whereby the focal presence of tau pathology in midbrain and basal ganglia led to widespread dysfunction in the functional connectivity across cortical networks. This is a generalisation of previous demonstrations of disruption of spatially distant connectivity networks resulting from damage to computational hubs within that network. For example, semantic dementia (semantic variant primary progressive aphasia) results from focal neurodegeneration of anterior temporal lobe. Anterior temporal lobe is proposed to perform a binding function for multi-modal semantic knowledge, which is itself distributed throughout cortex. Disruption of anterior temporal lobe function therefore leads to widespread abnormalities in functional connectivity, not only with temporal lobe itself but also between its neighbours (Guo et al., 2013). In PSP, I have demonstrated that the loss of inter-regional connectivity through subcortical structures results in greater cortico-cortical connectivity. This was at the cost of reduced network participation, eigenvector centrality and closeness centrality and increased betweenness centrality, together indicating that connectivity patterns were becoming more modular and information transfer was taking a less direct route. This goes some way towards
accounting for the phenotypic characteristics of subcortical cognitive impairment, including apathy and bradyphrenia.

Computational analyses are not without their pitfalls and the potential for misinterpretation. I minimised such risks by directly assessing focussed scientific hypotheses within a systems identification framework (Figure 5-1) (Goodwin and Payne, 1977, Ljung, 1998, Stokes and Purdon, 2017). Such a framework is equally applicable to the assessment of behavioural and neurophysiological data, and was at the core of my analysis strategies in chapters 2, 3 and 4. In this process it is vital to explicitly consider the assumptions made at every step, and to design the experiment in such a way that these assumptions can either be directly tested or rendered inconsequential. For example, in chapter 3 I formulated my candidate behavioural Bayesian inference models within a predictive coding framework, but they were equally applicable to any hierarchical generative model set in which top-down predictions are fed back to lower-order processing levels. I was then able to directly test this hypothesis (of bi-directional communication between higher and lower-order processing levels) with a second systems identification process based on neuronal coherence and Granger causality between left frontal and temporal regions of interest. In this way, I was able to provide direct evidence for causal top-down prediction signals, and to go on to precisely define the neural correlates and behavioural consequences of their disruption.
Figure 5-1 Systems identification framework
After Stokes and Purdon (2017). The process of making inference about brain function from computational modelling of behavioural and neurophysiological data.
5.2 Clinical implications of inflexible predictions in nfvPPA

The discussion in Chapter 3 was focussed on how investigations of the neurophysiological abnormalities of nfvPPA can provide insight into how the healthy brain processes language. In this section I change the focus, to address the question of how the neurophysiological abnormalities of nfvPPA that I have demonstrated can account for the clinical symptoms and signs of this disease. I begin (5.2.1) by considering the subjective speech comprehension symptoms that led to the initial conceptualisation of the study. I then go on (5.2.2) to address the basic auditory processing deficits that have previously been observed in nfvPPA, by first demonstrating these abnormalities in my cohort and then providing an explanation in terms of inflexible predictions. Finally (5.2.3), I place my results in the wider context of theories of grammatical comprehension. I explain how a synthesis of the results of chapters 3 and 4 can provide a perspective in which inflexible predictions can account for receptive agrammatism in nfvPPA.

5.2.1 Inflexible predictions can explain speech comprehension symptoms in nfvPPA

Chapter 3 was motivated by the observation that, in clinic, patients with nfvPPA frequently complain of difficulties in hearing speech that are disproportionate to any measured deficits and resistant to hearing aids. In some cases, this can be a presenting or prodromal feature of the illness, developing in advance of the output aphasia (Iizuka et al., 2007). Similarly, patients not meeting diagnostic criteria for nfvPPA but with a demonstrable metabolic or structural abnormality of inferior frontal brain regions have been reported to suffer from apperceptive auditory agnosias leading to amusia (Confavreux et al., 1992) or pure word deafness (Otsuki et al., 1998). Patients frequently describe using cross-modal cues such as lip reading or reading to support speech comprehension, even in quiet environments (Jörgens et al., 2008).

The symptom rating scales in my patient cohort (Figure 3-2) demonstrated a dissociation in subjective assessment between understanding speech in noise (which was rated as equally difficult by both patients and controls), and speech in quiet (which only patients rated as difficult). Greater difficulty in quiet environments may seem counterintuitive at first, but it is consistent with the primary abnormality being one of
inflexible predictions. Successful perception and comprehension of speech requires continuous updating of predictions based on sentential context and other cues. In a noisy environment, it is beneficial for listeners to rely heavily on these prior predictions as the patients do, because the sensory signal to noise ratio is poor. In a quiet environment, however, this is a suboptimal strategy as greater reliance can and should be placed on more precise or informative sensory inputs. If patients are unable to flexibly adapt the precision of their predictions to quiet listening environments, speech-in-quiet will remain difficult. It might be that globally strong predictions in nfvPPA are an adaptation to their inflexibility, as dysfunctioning networks are forced to choose one prediction strength for all scenarios. If, instead, predictions were globally weakened, this would be beneficial to the perception of speech-in-quiet but detrimental to speech-in-noise, perhaps having a greater overall cost to intelligibility in a dynamic environment (see (Wolpe et al., 2016) for similar arguments in motor control).

In the main experiment in chapter 3, I probed for subjective clarity ratings around 1050ms after the onset of the spoken word. It could be argued that the abnormal precision of prior expectations in patients might be adaptive to the experimental context. Delayed processing might mean that they have less time for predictions to be enacted before a decision must be made, and therefore stronger predictions are required if they are to have meaningful effects. In this experiment, I demonstrated that the neural processing of perceptual predictions was delayed in every patient compared to controls (Figure 3-11), and that in some individuals this processing continued beyond the analysis window. It could therefore be argued that my finding of increased prior precision in patients might have been attenuated if clarity ratings were requested after a longer delay. However, predictions on this slower timescale would be of limited real-world consequence for speech perception and comprehension, because content-containing words such as nouns and verbs are separated by similarly short intervals in typical sentences (Altmann and Kamide, 1999), the temporal range in which human perception is optimal (Cope et al., 2011, Cope et al., 2014a). Visual cues from lip reading also operate over millisecond timescales and are mediated by similar increases in fronto-temporal functional connectivity to those I demonstrate here (Giordano et al., 2017). Similarly, I included a delay of around 1050ms between the written word and the onset of the spoken word. For control participants and patients alike, this was
sufficient time to instantiate an appropriate prediction for upcoming speech signals (Figure 3-10).

5.2.2 Inflexible predictions can explain auditory processing abnormalities in nfvPPA

Goll et al. (2010) demonstrated that patients with nfvPPA were more impaired than those with Semantic Dementia at tasks assessing auditory perceptual function, despite similar disease duration and non-verbal cognitive impairment. The 12 nfvPPA patients in that study were rather more advanced than those I assessed in chapters 3 and 4, having an average Mini-Mental Status Examination (MMSE) score of 21/30, compared to an average MMSE of 28/30 in my cohort. The tasks at which patients with nfvPPA were most impaired compared to controls were a same/different judgment of two spectrally shaped harmonic series (essentially a test of timbre discrimination) and a same/different judgment of the semantic identity of two auditory stimuli. These deficits were less severe in analogous tasks undertaken in the visual domain. In a subsequent study containing 5 patients with nfvPPA, (Goll et al., 2011) demonstrated that the magnitude of impairment was similar across a range of auditory tasks in which the acoustic judgment was same/different or up/down, but that they performed better when asked to assess a property of the stimulus source such as big/small or tool/animal.

Grube et al. (2016) sought to quantify this impairment by using adaptive psychophysical methods to assess a comprehensive auditory processing battery in 18 consecutive patients with Primary Progressive Aphasia. This battery comprised 16 tasks, assessing sensitivity to changes in stimulus pitch, rhythm and temporal dynamics. All of the tasks required two-alternative forced choice same/different (either as AB or AxB comparisons), or change-detection judgments. They demonstrated that, on average, the six patients in their cohort with nfvPPA were more than 1 standard deviation poorer than control participants at all tasks, with an overall mean Z-score of approximately -1.9. This impairment was not present to the same extent in groups with semantic dementia (mean Z-score -0.7) or logopenic aphasia (mean Z-score -0.1). Performance was particularly poor where a comparison of prolonged stimulus sequences was required.
In my, larger, cohort of 11 nfvPPA patients described in chapter 3 I exactly replicated a subset of the tasks used by Grube et al. (2016). I selected a cross section of tasks on which the patients with nfvPPA had displayed particular difficulties, covering a range of auditory processing from simple to complex. These comprised pitch change detection (Grube task P1), 2Hz and 40Hz frequency modulation detection (Grube tasks M1 and M2), and dynamic ripple discrimination (Grube task M4).

**Basic Auditory Processing**

![Basic Auditory Processing](image)

Figure 5-2 Basic auditory processing in nfvPPA, after Grube et al. (2016). The tasks employed were pitch change detection, 2Hz FM detection, 40Hz FM detection, and dynamic ripple density discrimination. Statistical group differences were assessed with unpaired t-tests with unequal variance, and are indicated by * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

My results replicated the primary finding of Grube et al. (2016), that patients with nfvPPA perform very poorly at some tasks of basic auditory processing (Figure 5-2). This did not seem to have a trivial explanation like an inability to sustain attention or yes/no confusion, as the individual adaptive tracks had a similar shape in patients and controls, with consistent correct responses in ‘easy’ trials and, once threshold is reached, flat profiles maintained for the remainder of a run (Figure 5-3).
Average Adaptive Tracks by Group

Figure 5-3 Overall average adaptive tracks by group

The pattern of performance was highly variable between individuals, but highly consistent within individuals. As can be seen in the individual Z-scores for each task, some patients were able to consistently perform some discrimination tasks in the normal range, while being dozens of standard deviations poorer than the mean in other tasks. Patients who performed well on a particular task continued to perform well if it was repeated but, even after repeated practice, they remained unable to perform well on tasks that they had previously found difficult.

On the face of it, these consistent demonstrations of impaired basic auditory function in nfvPPA seem surprising as the Bayesian VBM provides evidence for no atrophy in primary auditory regions (Figure 3-1 E), and at post mortem patients with nfvPPA do not display disproportionate pathology in either primary auditory cortex or auditory brainstem nuclei. Further, it is known that patients with progressive supranuclear palsy, who do have severe brain stem atrophy and focal pathology that involves the inferior colliculi, continue to display complex auditory psychophysical effects late in disease (Hughes et al., 2014). Similarly, my patients with nfvPPA demonstrated good performance at identifying vocoded words, performing almost as well as controls (Figure 3-3 D). Finally, the pattern of psychophysical deficits was observed to be highly variable between individuals (Figure 5-4), but highly consistent within individuals;
impairment of a bottom-up perceptual process would predict a consistent profile of performance that might vary in severity, while what I observe is that all individuals perform very poorly on some tasks, but the relative difficulty of the tasks varies between individuals. This is fundamentally different from the performance profile seen in conditions where abnormal auditory processing could in itself be a sufficient explanation for patients’ symptomatology. For example, patients with Wernicke’s aphasia due to lesions of left temporo-parietal junction display normal low-level processing of pitch and timbre information, but impaired ability to detect or process spectral flux (those changes in the frequency domain that occur over time) (Robson et al., 2013). This is explicable as higher stages of the auditory system process progressively longer temporal windows (Cope et al., 2011). Further, it can potentially account for the receptive symptomatology in Wernicke’s aphasia by accounting for impaired mapping of auditory objects to phonemes as the neural architecture of planum temporale is disrupted (Griffiths and Warren, 2004).

Overall, therefore, a higher level, cortical explanation must be invoked to account for these psychophysical findings. This effect has previously been understood in terms of impaired working memory (Grube et al., 2016), but it is worth considering whether it might also be explained by the abnormalities of predictive coding demonstrated here,
and indeed whether the two explanations might in fact reflect the same underlying cognitive process.

Basic auditory processing is traditionally assessed with two- or three-alternative forced choice psychophysical paradigms, with the difference between exemplars adaptively modified to track a given performance percentile (Levitt, 1971). A predictive coding model can also be applied to this experimental context, in which subjects make a decision based on the location of the peak of their posterior in the perceptual dimension of interest. The distribution of this posterior is based on a prediction that is modified by prediction error induced by sensory input. For example, in a task where one is asked to detect the presence of a pitch change the subject might listen to the first pitch and then set up a prediction that the second pitch would be unchanged. A decision would be made by performing two-point discrimination on the peak locations of prior and posterior distributions. If all subjects establish similar predictions, the accuracy of this decision process is dependent only on the precision of the sensory input, as is the intent of adaptive measures in psychoacoustics (Levitt, 1971). If, however, subjects with nfVPPA make more precise predictions, the perceptual distance between prior and posterior would be reduced, leading to poor discriminatory performance even though the sensory input is unchanged. This also explains the lack of a hierarchical relationship in performance profiles (Figure 5-4); for example to discriminate the density of dynamic ripples it is necessary to be able to process frequency modulations, and yet some patients were able to perform within the normal range at discriminating ripples whilst seemingly unable to detect its building blocks (i.e. frequency modulation).

This explanation in terms of abnormally precise predictions is not exclusive of that in terms of impaired working memory (which could be modelled here as a drift in the location of the prior distribution over time), and indeed both processes could be occurring simultaneously. The argument I make is simply that it is possible that inflexible perceptual predictions, of themselves, are a sufficient explanation for measured impairments in basic sound discrimination in nfVPPA.
5.2.3 Inflexible predictions can explain receptive agrammatism in frontal aphasias

Agrammatism is a prominent symptom in both nfvPPA and its vascular analogue Broca-type aphasia. In ecological contexts, patients are observed to have particular difficulties understanding complex grammatical structures containing hierarchical structures or the passive voice (Goodman and Bates, 1997, Grodzinsky, 2000). At first glance this observation seems at odds with the empirical findings I presented in chapter 4, where I demonstrated patients to be equally impaired at the acquisition of grammatical rules based on linear transition probabilities, configurational position or hierarchical structure. However, these observations can be unified by inflexible perceptual predictions for sentential structure.

In most European languages, clauses are only disambiguated after the presentation of at least one noun. For example, when hearing one of the two sentences “The boy pushed the girl,” or “The boy was pushed by the girl,” the listener must either wait to hear the tensed verb before undertaking any sentential processing, or make a prediction about the likely relationship of ‘the boy’ to what follows. Sub-clausal descriptions are subject to similar constraints: the subject-relative sub-clause in, “The doctor that annoyed the nurse ran a clinic,” is more easily and quickly parsed than the object-relative sentence, “The doctor that the nurse annoyed ran a clinic,” (Ford, 1983). Linear, subject-relative sentences are much more frequently encountered across a range of European languages (Mak et al., 2002). It would therefore be reasonable to hold a prior expectation for these more frequent, subject-oriented linear word orders. Indeed, frequent exposure to non-linear sentence structures has been shown to increase their processing speed through statistical learning (Wells et al., 2009). There are a number of mechanistic explanations for this effect, but one view is formalised in the perspective-shifting account of syntactic processing, in which the listener must re-orient their mental imagery to an unexpected agent to parse a sentence or sub-clause (for a review, see (Traxler et al., 2002)). If patients with frontal aphasias are less able to flexibly modify their prior expectations and shift perspective on the basis of grammatical cues, it could account for their selective behavioural impairment.

Support for this view comes from the observation that sentence predictability modulates the difficulty of syntactic comprehension in healthy individuals. Inanimate
objects are more frequently the subject of object-relative descriptions across a range of languages (Mak et al., 2002). Consequently, when readers or listeners encounter inanimate objects as the first noun in a sentence, their perceptual prediction that the sentence will follow a linear progression will be less strong. This leads to the observation that non-linear structures are processed more quickly and require less re-reading (so-called ‘first-pass regressions’) when the first noun is inanimate, rather than animate (Traxler et al., 2002). Similarly, manipulating semantic predictability can reduce or eliminate syntactic complexity effects by ensuring that the sentence could only be legitimately interpreted in one way (Traxler et al., 2005). These two effects (animacy and semantic content) interact, such that higher-order understanding of the topic combines with implicit statistical learning to produce optimal perceptual predictions (Mak et al., 2006). When these predictions are violated during reading, electrophysiological correlates are observed with a peak at around 400ms (the N400 response) (Dambacher et al., 2006, Van Petten and Luka, 2012), a strikingly similar latency to that observed in my control participants in chapter 3 (Figure 3-11).

Thus, the inflexible predictions that I demonstrate in chapter 3 are able to reconcile the seeming inconsistency between the findings of chapter 4 and the observed symptomatology of agrammatic aphasia, by proposing that the core perceptual deficit occurs downstream of the syntactic cue. In other words, although the patients have not lost the ability to implicitly learn and understand more complex sequencing relationships, they are unable to use these subtle word-ordering or tense cues to perspective-shift and modify their initial prediction for a linear presentation. This view is consistent with emerging perspectives of linguistic processing as a specialised function of a more general cognitive computational system for complex and flexible thought, based on dynamic functional interactions between inferior frontal and superior temporal cortex (Friederici et al., 2017, Milne et al., 2018).
Summary

The dementias are persistent or progressive disorders affecting more than one cognitive domain that interfere with an individual’s ability to function at work or home, and represent a decline from a previous level of function. In this thesis I consider the neurophysiology of dementia at a number of levels. I investigate the ways in which the connectivity and function of the brain predisposes to the specific focal patterns of neurodegeneration seen in the various dementias. I aim to identify the mesoscopic changes that occur in individuals with neurodegeneration and how these relate to their cognitive difficulties. I show how, by assessing patients in whom there is focal disruption of brain networks and observing the outcomes in comparison to controls, I can gain insight into the mechanisms by which the normal brain makes predictions and processes language.

In Chapter 1, I set the scene for the focused experimental investigations of model diseases by beginning with an introductory, clinically-focused review that sets out the features, aetiology, management, epidemiology and prognosis of the dementias. This places these model diseases in the context of the broader clinical challenge posed by the dementias.

In Chapter 2, I turn to ‘prototypical’ model diseases that represent neurodegenerative tauopathies with predominantly cortical (Alzheimer’s disease, AD) and subcortical (Progressive Supranuclear Palsy, PSP) disease burdens. I investigate the neurophysiological causes and consequences of Tau accumulation by combining graph theoretical analyses of resting state functional MR imaging and in vivo ‘Tau’ PET imaging using the ligand AV-1451. By relating Tau distribution to the functional connectome I provide in vivo evidence consistent with ‘prion-like’ trans-neuronal spread of Tau in AD but not PSP. This provides important validation of disease modification strategies that aim to halt or slow down the progression of AD by sequestration of pathological Tau in the synapse. In contrast, I demonstrate associations consistent with regional vulnerability to Tau accumulation due to metabolic demand and a lack of trophic support in PSP but not AD. With a cross-sectional approach, using Tau burden as a surrogate marker of disease severity, I then go on to show how the changes in functional connectivity that occur as disease
progresses account for the contrasting cognitive phenotypes in AD and PSP. In advancing AD, functional connectivity across the whole brain becomes increasingly random and disorganised, accounting for symptomatology across multiple cognitive domains. In advancing PSP, by contrast, disrupted cortico-subcortical and cortico-brainstem interactions meant that information transfer passed through a larger number of cortical nodes, reducing closeness centrality and eigenvector centrality, while increasing weighted degree, clustering, betweenness centrality and local efficiency. Together, this resulted in increasingly modular processing with inter-network communication taking less direct paths, accounting for the bradyphrenia characteristic of the ‘subcortical dementias’.

From chapter 3 onwards, I turn to the in-depth study of a model disease called non-fluent variant Primary Progressive Aphasia (nfvPPA). This disease has a clear clinical phenotype of speech apraxia and agrammatism, associated with a focal pattern of mild atrophy in frontal lobes. Importantly, general cognition is usually well preserved until late disease.

In chapter 3 itself, I relate an experiment in which patients with nfvPPA and matched controls performed a receptive language task while having their brain activity recorded with magnetoencephalography. I manipulated expectations and sensory detail to explore the role of top-down frontal contributions to predictive processes in speech perception. I demonstrate that frontal neurodegeneration led to inflexible and excessively precise predictions, and that fronto-temporal interactions play a causal role in reconciling prior predictions with degraded sensory signals. The discussion here concentrates on the insights provided by neurodegenerative disease into the normal function of the brain in processing language. Overall, I demonstrate that higher level frontal mechanisms for cognitive and behavioural flexibility make a critical functional contribution to the hierarchical generative models underlying speech perception.

In chapter 4, I precisely define the sequence processing and statistical learning abilities of patients with nfvPPA in comparison to patients with non-fluent aphasia due to stroke and neurological controls. I do this by exposing participants to a novel, mixed-complexity artificial grammar designed to assess processing of increasingly complex sequencing relationships, and then assessing the degree of implicit rule learning. I demonstrate that agrammatic aphasics of two different aetiologies are not
disproportionately impaired on complex sequencing relationships, and that the learning of phonological and non-linguistic sequences occurs independently in health and disease.

In chapter 5, I summarise the synergies between the experimental chapters, and explain how I have applied a systems identification framework to a diverse set of experimental methods, with the common goal of defining the physiology of dementia. I then return to the results of chapter 3 with a clinical focus to explain how inflexible predictions can account for subjective speech comprehension difficulties, auditory processing abnormalities and (in synthesis with chapter 4) receptive agrammatism in nfvPPA.

Overall, this body of work has contributed to knowledge in several ways. It has achieved its tripartite aims by:

1) Providing *in vivo* evidence consistent with theoretical models of trans-neuronal Tau spread (chapter 2), and a comprehensive clinical account of the previously poorly-understood receptive symptomatology of nfvPPA (chapter 5), thus demonstrating that systems neuroscience can provide a translational bridge between the molecular biology of dementia and clinical trials of therapies and medications. In this way, I begin to disentangle the network-level *causes* of neurodegeneration from its *consequences*.

2) Providing evidence for a causal role for fronto-temporal interactions in language processing (chapter 3), and demonstrating domain separation of statistical learning between linguistic and non-linguistic sequences (chapter 4), thus demonstrating that studies of patients with neurodegenerative disease can further our understanding of normative brain function.

3) Successfully integrating neuropsychology, behavioural psychophysics, functional MRI, structural MRI, magnetoencephalography and computational modelling to provide comprehensive research training, as the platform for a future research programme in the physiology of dementia.
References


Lebert F. Behavioral benefits of trazodone are sustained for the long term in frontotemporal dementia. 2006.


McQueen JM, Norris D, Cutler A. Are there really interactive processes in speech perception? Trends in Cognitive Sciences. 2006;10(12):533-.


Rogalski EJ, Saxon M, McKenna H, Wieneke C, Rademaker A, Corden ME, Borio K, Mesulam M-M, Khayum B. Communication Bridge: A pilot feasibility study of Internet-


Schwartz J-L, Savariaux C. No, there is no 150 ms lead of visual speech on auditory speech, but a range of audiovisual asynchronies varying from small audio lead to large audio lag. PLoS Comput Biol. 2014;10(7):e1003743.


Stokes PA, Purdon PL. A study of problems encountered in Granger causality analysis from a neuroscience perspective. Proc Natl Acad Sci U S A. 2017 Aug 22;114(34):E7063-E72.


von Helmholtz H. Helmholtz's treatise on physiological optics: Optical Society of America; 1925.


Woollams AM, Ralph MAL, Plaut DC, Patterson K. SD-squared: on the association between semantic dementia and surface dyslexia. Psychol Rev. 2007;114(2):316.


Appendix: Published work resulting from this thesis

At the time of binding, chapters 2, 3 and 4 have resulted in journal publications. These publications now follow in their final forms as appendices to this thesis.

Chapter 2 forms the basis of a paper that has been published in *Brain* (*Cope et al.*, 2018).

Chapter 3 forms the basis of a paper that has been published in *Nature Communications* (*Cope et al.*, 2017a).

Chapter 4 forms the basis of a paper that has been published in *Neuropsychologia* (*Cope et al.*, 2017b).
Tau burden and the functional connectome in Alzheimer’s disease and progressive supranuclear palsy

Thomas E. Cope,1 Timothy Rittman,1 Robin J. Borchert,1 P. Simon Jones,1 Deniz Vatansever,1,2,3,4 Kieren Allinson,5 Luca Passamonti,1 Patricia Vazquez Rodriguez,1 W. Richard Bevan-Jones,1,4 John T. O’Brien4,* and James B. Rowe 1,6,*

Alzheimer’s disease and progressive supranuclear palsy (PSP) represent neurodegenerative tauopathies with predominantly cortical versus subcortical disease burden. In Alzheimer’s disease, neuropathology and atrophy preferentially affect ‘hub’ brain regions that are densely connected. It was unclear whether hubs are differentially affected by neurodegeneration because they are more likely to receive pathological proteins that propagate trans-neuronally, in a prion-like manner, or whether they are selectively vulnerable due to a lack of local trophic factors, higher metabolic demands, or differential gene expression. We assessed the relationship between tau burden and brain functional connectivity, by combining in vivo PET imaging using the ligand AV-1451, and graph theoretic measures of resting state functional MRI in 17 patients with Alzheimer’s disease, 17 patients with PSP, and 12 controls. Strongly connected nodes displayed more tau pathology in Alzheimer’s disease, independently of intrinsic connectivity network, validating the predictions of theories of trans-neuronal spread but not supporting a role for metabolic demands or deficient trophic support in tau accumulation. This was not a compensatory phenomenon, as the functional consequence of increasing tau burden in Alzheimer’s disease was a progressive weakening of the connectivity of these same nodes, reducing weighted degree and local efficiency and resulting in weaker ‘small-world’ properties. Conversely, in PSP, unlike in Alzheimer’s disease, those nodes that accrued pathological tau were those that displayed graph metric properties associated with increased metabolic demand and a lack of trophic support rather than strong functional connectivity. Together, these findings go some way towards explaining why Alzheimer’s disease affects large scale connectivity networks throughout cortex while neuropathology in PSP is concentrated in a small number of subcortical structures. Further, we demonstrate that in PSP increasing tau burden in midbrain and deep nuclei was associated with strengthened cortico-cortical functional connectivity. Disrupted cortico-subcortical and cortico-brainstem interactions meant that information transfer took less direct paths, passing through a larger number of cortical nodes, reducing closeness centrality and eigenvector centrality in PSP, while increasing weighted degree, clustering, betweenness centrality and local efficiency. Our results have wide-ranging implications, from the validation of models of tau trafficking in humans to understanding the relationship between regional tau burden and brain functional reorganization.

1 Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK
2 Department of Psychology, University of York, York, UK
3 Division of Anaesthesia, Department of Medicine, University of Cambridge, Cambridge, UK
4 Department of Psychiatry, University of Cambridge, Cambridge, UK
5 Department of Pathology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
6 Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK
Introduction

Alzheimer’s disease and progressive supranuclear palsy (PSP) are both characterized by intracellular neurofibrillary lesions containing hyper-phosphorylated filamentous tau inclusions (Goedert and Spillantini, 2006). However, the diseases differ in: (i) the distribution of these tau inclusions; (ii) the balance of expression of tau isoforms; and (iii) the ultrastructure of tau filaments. Here, we test the impact of these differences in tau pathology on the reorganization of large-scale functional brain connectivity architecture.

Alzheimer’s disease is characterized by widespread extracellular deposition of amyloid-β and paired helical filaments of tau with three (3R) and four (4R) repeats in the microtubule-binding domain (Sisodia et al., 1990; Liu et al., 2001). In Alzheimer’s disease, these pathological proteins arise early in the transentorhinal cortex, from where they spread to limbic regions, followed by inferior frontal and parietal cortex (Braak and Braak, 1991, 1995). Although tau and amyloid-β have complex and synergistic effects (Nisbet et al., 2015), there is converging evidence that tau mediates direct toxic effects on neurons and synaptic plasticity (Ballatore et al., 2007; Roberson et al., 2007; Ittner and Götz, 2011; Myeku et al., 2016) and correlates with hypometabolism and symptomatology in Alzheimer’s disease (Lehmann et al., 2013; Ossenkoppele et al., 2015, 2016). By contrast, pathological tau deposits in PSP are composed of straight filaments of predominantly 4R tau (Taniguchi-Watanabe et al., 2016). This is most prominent in midbrain and deep brain nuclei early in the course of the Richardson’s syndrome variant of PSP, spreading to cortical regions in advanced stages of the disease (Williams et al., 2007).

To estimate the burden of tau pathology in vivo, PET ligands have been developed, including 18F-AV-1451 (Chien et al., 2013; Xia et al., 2013; Ossenkoppele et al., 2015, 2016). While there has been significant debate about the specificity of this ligand for tau, especially in the non-Alzheimer’s dementias (Bevan-Jones et al., 2017a; Xia and Dickerson, 2017), the distribution of 18F-AV-1451 binding recapitulates Braak staging in Alzheimer’s disease (Schwarz et al., 2016) and correlates with post-mortem neuropathology in primary tauopathies (Smith et al., 2016; Passamonti et al., 2017). 18F-AV-1451 binding has been shown to correlate with cognitive performance in Alzheimer’s disease (Johnson et al., 2016) and healthy older adults (Schöll et al., 2016) more robustly than quantitative amyloid-β imaging (Brier et al., 2016). The 18F-AV-1451 ligand also effectively distinguishes between Alzheimer’s disease and PSP cases based on both the intensity and regional distribution of its binding potential (BPND) (Passamonti et al., 2017).

In many neurodegenerative disorders, neuropathology and atrophy are most marked in those brain regions that are densely connected, both at the structural (Crossley et al., 2014) and functional level (Dai et al., 2014). In graph-theoretical terms, these densely connected regions are usually referred to as ‘hubs’ (Buckner et al., 2009). There are a number of hypotheses as to why hubs are vulnerable to neurodegeneration. First, pathological proteins may propagate trans-neuronally, in a prion-like manner (Prusiner, 1984; Baker et al., 1994; Goedert, 2015) such that highly connected regions are more likely to receive pathology from ‘seed’ regions affected in early stages of the disease (Zhou et al., 2012), leading to neurodegeneration that mirrors structural and functional brain connectivity (Raj et al., 2012, 2015; Abdelnour et al., 2014). Alternatively, hubs might be selectively vulnerable to a given level of pathology, due to a lack of local trophic factors (Appel, 1981), higher metabolic demands (Saxena and Caroni, 2011; de Haan et al., 2012), or differential gene expression (Rittman et al., 2016).

These alternate hypotheses lead to different predictions about the relationship between tau burden and connectivity. The trans-neuronal spread hypothesis predicts that regions that are more strongly interconnected would accrue more tau pathology. This would manifest as higher tau burden in nodes with larger weighted degree, which is a measure of the number and strength of functional connections involving each node. In contrast, if hubs are vulnerable to tau accumulation because of increased metabolic demand this might manifest as a positive relationship between tau burden and participation coefficient, which is a measure of the proportion of a node’s connections that are with other neural communities, and is the graph metric that is most closely correlated with metabolic activity (Chennu et al., 2017). Finally, if trophic support is an important factor in tau accumulation, this might manifest as a negative relationship between tau burden and clustering coefficient; nodes with less tightly clustered connectivity patterns might have more vulnerable trophic supply.

Here we go beyond previous associative studies to examine, in the same subjects, the relationship between in vivo tau burden, as measured by the PET ligand 18F-AV-1451,

Keywords: Alzheimer’s disease; progressive supranuclear palsy; tau; functional connectivity; graph theory

Abbreviations: BPND = binding potential of PET ligand; MCI = mild cognitive impairment; PSP = progressive supranuclear palsy

E-mail: thomascope@gmail.com

Herchel Smith Building, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK

Correspondence to: Dr Thomas E. Cope, Herchel Smith Building, Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK

E-mail: thomascope@gmail.com
BP\textsubscript{ND}, and functional connectivity, as summarized by graph theoretic measures based on resting state (task-free) functional MRI. We test the following linked hypotheses:

(i) Brain regions that are normally more densely interconnected accrue more tau pathology.

(ii) In Alzheimer’s disease, where tau accumulation is predominantly cortical, the functional consequence is that affected nodes become more weakly connected and local efficiency of information transfer is reduced.

(iii) In PSP-Richardson’s syndrome, where neurodegeneration associated with tau accumulation is most severe in midbrain and basal ganglia, the functional consequence of disrupted cortico-subcortical and cortico-brainstem interactions (Gardner et al., 2013) is that indirect cortico-cortical connections become stronger.

Materials and methods

Participants

All patients had mental capacity to take part in the study and provided informed consent. Study procedures were approved by the National Research Ethics Service. We recruited 17 patients with Alzheimer’s disease, as evidenced by a clinical diagnosis of probable Alzheimer’s dementia according to consensus criteria ($n = 10$) (McKhann et al., 2011), or a clinical diagnosis of mild cognitive impairment (MCI) and a positive amyloid PET scan ($n = 7$) (Klunk et al., 2004; Okello et al., 2009). Patients with MCI and evidence of amyloid-$\beta$ were included to ensure the largest possible variability in tau burden within the Alzheimer’s disease group. We also recruited 17 patients with PSP-Richardson’s syndrome by 1996 criteria (Litvan et al., 1996). Retrospective case review confirmed that all PSP subjects also met the revised 2017 criteria for PSP-RS (Höglinger et al., 2017). Twelve age-matched controls were also examined. Participant demographics are given in Table 1 and detailed neuropsychological test results for each subject are shown in Supplementary Table 1.

Study procedures

This study formed part of the NIMROD (Neuroimaging of Inflammation in Memory and Related Other Disorders) project, for which the trial protocol containing general methods has been previously published (Bevan-Jones et al., 2017).

All participants underwent the Addenbrooke’s Cognitive Examination-Revised (ACE-R), 11 min of resting state functional MRI at 3T, and $^{11}$F-AV-1451 PET imaging. Participants with PSP were also examined according to the PSP Rating Scale (Golbe and Ohman-Strickland, 2007). Those with a clinical diagnosis of MCI had $^{11}$C-PiB PET imaging on a separate occasion—only those with increased uptake indicative of underlying Alzheimer’s pathology are reported here.

Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Age, years</th>
<th>Years of education</th>
<th>MMSE</th>
<th>ACE-R</th>
<th>PSP-RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s disease</td>
<td>17</td>
<td>71 (9)</td>
<td>14 (3)</td>
<td>25 (4)</td>
<td>72 (14)</td>
<td>-</td>
</tr>
<tr>
<td>PSP</td>
<td>17</td>
<td>69 (6)</td>
<td>12 (2)</td>
<td>27 (4)</td>
<td>83 (14)</td>
<td>41 (14)</td>
</tr>
<tr>
<td>Controls</td>
<td>12</td>
<td>67 (8)</td>
<td>16 (2)</td>
<td>29 (1)</td>
<td>96 (3)</td>
<td>-</td>
</tr>
</tbody>
</table>

Values are mean (SD). ACE-R = Addenbrooke’s Cognitive Examination, Revised Edition; MMSE = Mini-Mental State Examination; PSP-RS = Progressive Supranuclear Palsy Rating Scale.

MRI data acquisition and preprocessing

MRI was performed at the Wolfson Brain Imaging Centre, University of Cambridge, UK using a 3T Siemens Magnetom Tim Trio scanner with a Siemens 32-channel phased-array head coil (Siemens Healthcare).

A T$_1$-weighted magnetization-prepared rapid gradient-echo (MPRAGE) image was acquired with repetition time = 2300 ms, echo time = 2.98 ms, matrix = 236 $\times$ 240, in-plane resolution of 1 $\times$ 1 mm, 176 slices of 1 mm thickness, inversion time = 900 ms and flip angle = 9°.

Eyes-closed resting state (task-free) multi-echo functional imaging was carried out for 11 min. A total of 269 EPI image volumes were acquired with repetition time = 2430 ms, echo times = 13.00, 30.55 and 48.10 ms, matrix = 64 $\times$ 64, in-plane resolution of 3.75 $\times$ 3.75 mm, 34 slices of 3.8 mm thickness with an interslice gap of 0.38 mm, GRAPPA parallel imaging with an acceleration factor of 2 and bandwidth = 2368 Hz/pixel. The first six volumes were discarded to eliminate saturation effects and achieve steady state magnetization. Preprocessing employed the ME-ICA pipeline (https://wiki.cam.ac.uk/bmuwiki/MEICA) (Kundu et al., 2012, 2013), which uses independent component analysis to classify BOLD and non-BOLD signals based on the identification of linearly dependent and independent echo time dependent components. This provides an optimal approach to correct for movement-related and non-neuronal signals, and is therefore particularly well suited to our study, in which systematic differences in movement or head position might reasonably have been expected between patient groups. In fact, perhaps surprisingly, such differences were not observed (Supplementary Fig. 1)—between group ANOVAs demonstrated no significant differences between groups in terms of frame displacement before ($P = 0.88$) or after ($P = 0.42$) ME-ICA preprocessing, nor in terms of DVARS (Power et al., 2012) before ($P = 0.89$) or after ($P = 0.67$) ME-ICA preprocessing. Note that both movement parameters were approximately an order of magnitude lower after ME-ICA preprocessing.

The MPRAGE images were processed into the standard space using DARTEL (Ashburner, 2007), producing a study-specific template in stereotactic space. Each functional MRI

Table 1: Participant demographics

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Age, years</th>
<th>Years of education</th>
<th>MMSE</th>
<th>ACE-R</th>
<th>PSP-RS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s disease</td>
<td>17</td>
<td>71 (9)</td>
<td>14 (3)</td>
<td>25 (4)</td>
<td>72 (14)</td>
<td>-</td>
</tr>
<tr>
<td>PSP</td>
<td>17</td>
<td>69 (6)</td>
<td>12 (2)</td>
<td>27 (4)</td>
<td>83 (14)</td>
<td>41 (14)</td>
</tr>
<tr>
<td>Controls</td>
<td>12</td>
<td>67 (8)</td>
<td>16 (2)</td>
<td>29 (1)</td>
<td>96 (3)</td>
<td>-</td>
</tr>
</tbody>
</table>

Values are mean (SD). ACE-R = Addenbrooke’s Cognitive Examination, Revised Edition; MMSE = Mini-Mental State Examination; PSP-RS = Progressive Supranuclear Palsy Rating Scale.
series mean image was co-registered to the corresponding MPRAGE image. The whole functional MRI series was warped to the template space using the DARTEL flow fields.

To perform a whole brain graph-theoretical analysis that included the brainstem, cerebellum and subcortical structures, the Harvard-Oxford Cortical atlas and Harvard-Oxford Subcortical atlas, each thresholded at 25%, were combined. Additionally, a Freesurfer 6 brainstem parcellation of the MNI152 (2009 asymmetric) brain together with remaining Ventral DC completed the whole brain labelling. This atlas was sub-parcelled into 598 regions of approximately equal volume [mean 1.995, standard deviation (SD) 0.323 ml] such that each sub- parcel could be uniquely identified with an atlas region. The MNI-space parcellation was matched to the group standard space using inverse deformations following application of the ‘Population to ICBM Registration’ SPM function to the group template, with nearest neighbour interpolation.

The blood oxygen level-dependent (BOLD) time series for each node was extracted using the CONN functional connectivity toolbox (Whitfield-Gabrieli and Nieto-Castanon, 2012). Between-node association matrices were generated, and then z-transformed for further analysis.

**Functional connectivity assessment**

Graph theoretical analysis was used to investigate the global and local characteristics of brain networks. These metrics were calculated in python using the Maybrain software (github.com/rittmann/maybrain) and networkx (version 1.11). Reported metrics were calculated in each subject at absolute network density thresholds of 1–10% in 1% increments using a minimum spanning tree to ensure complete connectivity of the graph (Alexander-Bloch et al., 2010). Primary statistical analysis was performed at an intermediate density of 6%, with confirmatory analyses separately performed at all other densities. Weighted degree and participation coefficient were analysed in their raw forms, and all other metrics were dissociated from variation in degree by binarization after thresholding and normalization against 1000 random graphs with the same number of connections at each node.

The graph metrics assessed were:

(i) Weighted degree: the number and strength of functional connections involving each node.

(ii) Weighted participation coefficient: the proportion of a node’s functional connectivity that involves other nodes that are not part of its own community structure, as defined by the Louvain community detection algorithm. Nodes with high degree and low participation coefficient are ‘provincial hubs’ (i.e. display strong connectivity only within their own community), while those with high participation coefficient are connector nodes (Joyce et al., 2010).

(iii) Betweenness centrality: the number of shortest paths between any other two nodes that pass through the node of interest. Nodes that are important for the transfer of information between other nodes have high betweenness centrality.

(iv) Closeness centrality: the inverse of the path length between a node and all other nodes in the graph. This is the node-wise equivalent of global efficiency, which is the inverse sum of all the shortest path lengths in the graph.

(v) Local efficiency: the number of strong connections a node has with its neighbouring nodes. This reflects the robustness of local networks to disruption.

(vi) Eigenvector centrality: this measure quantifies the functional influence of a node on every other node in the graph, by weighting the importance of each nodal connection based on the influence of the nodes with which they connect.

(vii) Clustering coefficient: the fraction of triangular connections formed by a node with other nodes. In other words, a node is strongly clustered if a large proportion of its neighbours are neighbours of each other.

Nodal connectivity strength was assessed for comparison to weighted degree, to ensure that our results did not result from bias introduced by proportionate thresholding. This metric is related to weighted degree, but includes information from all strengths of connection between every pair of nodes. As such, it is more subject to functional MRI signal-to-noise ratio limitations, and it is not a suitable metric for whole-brain, cross-sectional analysis across individuals, but it can be used to make a node-wise, group average assessment analogous to that for weighted degree.

**Tau burden assessment**

Other than the atlas used to define regions of interest, all AV-1451 and PiB data acquisition and preprocessing steps were identical to those reported by Passamonti et al. (2017). Ligand preparation was carried out at the Wolfson Brain Imaging Centre, University of Cambridge, with high radiochemical purity. 18F-AV-1451 was produced with a specific activity of 216 ± 60 GBq/μmol at the end of synthesis (60 min). 11C-PiB specific activity was > 150 GBq/μmol. PET imaging was performed on a GE Advance PET scanner with a 15 min 68Ge transmission scan used for attenuation correction acquiring 58 frames of increasing duration. PET scanning was performed in 3D mode (63 image planes; 15.2 cm axial field of view; 5.6 mm transaxial resolution and 2.4 mm slice interval) 80–100 min after a 9.0 to 11.0 mCi bolus injection in frames of 4 × 5 min. Each 11C-PiB scan was acquired using a 8.5 to 15 mCi bolus injection immediately followed by a 60-min dynamic acquisition in 69 frames (12 × 15 s, 57 × 60 s). Scans were reconstructed on the GE Advance scanner using the PROMIS 3D filtered back projection algorithm, correcting for randoms, dead time, normalization, sensitivity and scatter attenuation.

The cerebellar grey matter was used as a reference tissue to express the distribution volume ratio (DVR) for the 11C-PiB PET data for each MCI participant. PiB scans were classified as positive if the average SUVR values across the cortex were more than 1.5 times that of the cerebellar regions of interest. Seven MCI participants met this criterion and were included in the study.

The non-displaceable binding potential of 18F-AV-1451 was assessed at each region of interest in the sub-parcellated atlas after rigid registration of each subject’s dynamic PET image series to their T1-weighted MRI scan. This was normalized against the superior cerebellum, a reference region considered to have no tau pathology in either PSP or Alzheimer’s disease (Williams et al., 2007; Schöll et al., 2016; Passamonti et al., 2017). These data were corrected for white matter and CSF partial volumes by calculating the ordinary least squared
solution for the binding potential map voxel-wise in each region for grey plus white matter segments, each smoothed to PET resolution.

**Statistical approach**

A stereotyped statistical approach was taken to the analysis of graph metrics of theoretical interest. All statistical analyses were performed in MATLAB 2015b (The Mathworks Inc., 2015), with the exception of moderation analysis, which was performed in R (R Core Team, 2016).

A group-averaged analysis assessed the first hypothesis by examining the node-wise relationships between \(^{18}\text{F}-\text{AV-1451}\) binding potential and weighted degree, weighted participation coefficient, and clustering coefficient. These analyses were performed across the whole brain and within 10 intrinsic connectivity networks. To avoid circularity, where the regions of interest are defined by the test data, the intrinsic connectivity networks were independently derived from a publicly available dataset (Smith et al., 2009). A binary mask was constructed for each resting state network by thresholding at \(Z \geq 2.6 (P < 0.005\) uncorrected). Nodes were defined as belonging to a network if their centre of mass was within 5 mm of any positive voxel.

A between-subject analysis assessed the second and third hypotheses. This was undertaken at a variety of spatial scales. We first looked across the whole brain. For each individual, we first calculated a measure of disease-related tau burden. For participants with Alzheimer’s disease, in whom tau deposition increases in both magnitude and distribution as disease progresses (Braak and Braak, 1995), tau burden was calculated as average \(^{18}\text{F}-\text{AV-1451}\) binding potential across the whole brain. In PSP, tau deposition remains confined to brainstem and deep nuclei even in late disease (Williams et al., 2007); tau burden was therefore calculated as average \(^{18}\text{F}-\text{AV-1451}\) binding potential across the left and right thalamus, caudate, putamen, pallidum, accumbens, ventral diencephalon, midbrain, pons and medulla. Both methods were separately assessed for controls. These measures of individual tau burden were then correlated with whole-brain averaged graph metrics to assess the relationship between the metric in question and disease burden in each group separately. Moderation analysis was performed for each metric to assess whether the relationships differed between Alzheimer’s disease and PSP. It has recently been reported that patients with early onset Alzheimer’s disease have a higher tau burden than those with later onset, especially in later Braak stage regions (Scholl et al., 2017). Non-significant trend relationships between overall tau burden and age were observed in our cohort (Alzheimer’s disease \(r = -0.27\), PSP \(r = -0.02\), Control \(r = -0.38\); Supplementary Fig. 2). Cross-sectional analyses were therefore performed twice, with and without partialling out the effect of age on tau burden.

Regionality of the demonstrated effects was assessed within groups by correlating tau burden with the single subject graph metric values at each node; the gradient of the best fit linear regression within each group was the outcome measure. The nodal gradients in cerebellar regions were discarded (as this was the reference region for PET imaging), and the remaining regional maps were then collapsed into a vector. This was correlated with a matching vector of local tau burden at each node, calculated as the group-averaged increase in \(^{18}\text{F}-\text{AV-1451}\) binding potential. This resulted in a measure of the relationship between the distribution of disease-related change in each graph metric and regional deposition of tau.

**Results**

**Resting state connectivity differences exist between groups**

As a prelude to the detailed analysis of network features, we first confirmed that our groups differed in their resting state connectivity, by testing for differences between the group averaged association matrices (Fig. 1, top). Jenrich tests confirmed highly significant statistical differences between controls and Alzheimer’s disease (\(\chi^2 = 1.15 \times 10^8\), \(n_1 = 12\), \(n_2 = 17\), \(P < 0.0001\)); controls and PSP (\(\chi^2 = 975,547\), \(n_1 = 12\), \(n_2 = 17\), \(P < 0.0001\)); and Alzheimer’s disease and PSP (\(\chi^2 = 733,030\), \(n_1 = 17\), \(n_2 = 17\), \(P < 0.0001\)).

The nature of these differences would be opaque without abstraction techniques. Pairwise subtraction of the Fisher transformed association matrices (Fig. 1, bottom) revealed the presence of structure within the data, with both within-lobe and between-lobe group differences. While one can make general observations from these raw data, such as a prominent reduction in cerebello-cerebellar and brainstem-cerebellar connectivity in PSP, and a general reduction in between-region connectivity in Alzheimer’s disease, the insights would be limited. Instead, graph theoretic measures allow one to examine which specific properties of the network underpin disease-related differences. For this approach each region in the brain becomes a node in a graph, which is functionally connected to all other nodes with a finite strength that is given by their pairwise correlation over time. Each individual’s graph is then thresholded so that the highest ‘\(x\)’ percentage of connections survive. Effects are sought that are consistently present at a variety of network thresholds. Here we examine network thresholds from \(1 < x < 10\%\), representing a range of graphs from sparse to dense. Very sparse graphs contain less information and can miss important relationships. Conversely, very dense graphs are more subject to noise and, when binarized, begin to provide less meaningful information. Therefore, in what follows, we present the primary statistical analyses at an intermediate density of \(6\%\), with statistical detail given for this density. To assess the robustness of effects to different thresholding decisions we note separately the range of network densities exhibiting statistical significance.

**Findings in Alzheimer’s disease but not PSP are consistent with trans-neuronal spread of tau**

We hypothesized that trans-neuronal spread of tau would manifest as a positive relationship between \(^{18}\text{F}-\text{AV-1451}\) and weighted degree. As we have parcellated the brain into nodes of equal size, weighted degree is a measure of the...
volume of cortex to which a node is connected, and the strength of these connections. In Alzheimer’s disease, a strong positive correlation was observed at all network thresholds, such that the most strongly connected nodes had higher $^{18}$F-AV-1451 $\text{BP}_{\text{ND}}$ (Pearson’s $r = 0.48$, $P < 0.0001$, Spearman’s rho = 0.48, $P < 0.0001$) (Fig. 2A). A consistent relationship was not present in PSP (Pearson’s $r = -0.09$, Spearman’s rho = 0.12) (Fig. 2B) or controls (Pearson’s $r = 0.03$, Spearman’s rho = 0.11) (Fig. 2C). These patterns were present at all examined network density thresholds.

It is important to acknowledge that the degree of a node as defined by correlated activity is impacted by the size of the connectivity network to which it belongs (Power et al., 2013). Therefore, one concern might be that correlations between $^{18}$F-AV-1451 and degree are driven by the coincidence that Alzheimer’s disease happens to affect the default mode network (DMN), which happens to be a large intrinsic connectivity network. To exclude this possibility, we undertook two additional analyses. First, we re-examined $^{18}$F-AV-1451 and degree when nodes belonging to the DMN were excluded; this did not abolish the strong relationship (Pearson’s $r = 0.43$, $P < 0.0001$). Second, we examined the relationship within each intrinsic connectivity network (Fig. 3). A strong positive relationship between $^{18}$F-AV-1451 and degree was seen in 9 of 10 networks examined (Pearson’s $r \geq 0.4$, $P < 0.0001$). In contrast, no strong positive correlations observed in any network in PSP (Supplementary Fig. 3).

Finally, in each group we examined the relationship between each node’s unthresholded connectivity strength and $^{18}$F-AV-1451 $\text{BP}_{\text{ND}}$ (Supplementary Fig. 4). This is a local measure of the strength of functional connectivity that does not rely on thresholding the graph and therefore takes into account both strong and weak connections. The pattern of results observed in weighted degree was replicated; in Alzheimer’s disease, a positive correlation was observed ($r = 0.28$, $P < 0.0001$), but no significant relationship was observed in PSP ($r = -0.05$, $P = 0.22$) or controls ($r = 0.04$, $P = 0.31$).

**Findings in PSP but not Alzheimer’s disease are consistent with selective vulnerability**

We hypothesized that tau accumulation due to metabolic demand would manifest as a positive relationship between $^{18}$F-AV-1451 and weighted participation coefficient. This was not observed in Alzheimer’s disease; in fact there was a weak negative correlation (Pearson’s $r = -0.18$) (Fig. 4A).

In PSP, however, those nodes that displayed elevated $^{18}$F-AV-1451 were those that had the highest participation coefficient (Fig. 4B).
Similarly, we hypothesized that tau accumulation due to a lack of trophic support would manifest as a negative relationship between \( ^{18}\text{F-AV-1451} \) binding potential and clustering coefficient. In Alzheimer’s disease, the opposite relationship was observed; strongly clustered nodes were more likely to display elevated \(^{18}\text{F-AV-1451} \) binding (Pearson’s \( r = 0.33 \), \( P < 0.0001 \), Spearman’s rho = 0.48, \( P < 0.0001 \)), and the corresponding regression line is plotted for this group. (D) Between-subjects analysis of the relationship between global tau burden and each weighted degree at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden—for simplicity only their whole-brain average points are illustrated here. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (Alzheimer’s disease and PSP) was statistically significant. (E) The magnitude of disease-related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups. (F) Average \(^{18}\text{F-AV-1451} \) binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. (G) The local tau burden-related change in weighted degree is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. (H) Average raw values for weighted degree within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.

**Figure 2. Weighted degree.** (A–C) Group-averaged connection strength at each node, quantified by weighted degree, plotted against \(^{18}\text{F-AV-1451} \) binding potential at that node. A statistically significant linear relationship was demonstrated only in Alzheimer’s disease (AD) (Pearson’s \( r = 0.48 \), \( P < 0.0001 \), Spearman’s rho = 0.48, \( P < 0.0001 \)), and the corresponding regression line is plotted for this group. (D) Between-subjects analysis of the relationship between global tau burden and each weighted degree at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden—for simplicity only their whole-brain average points are illustrated here. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (Alzheimer’s disease and PSP) was statistically significant. (E) The magnitude of disease-related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups. (F) Average \(^{18}\text{F-AV-1451} \) binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. (G) The local tau burden-related change in weighted degree is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. (H) Average raw values for weighted degree within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.

**Alzheimer’s disease and PSP have opposite effects on the strength of cortical functional connectivity**

We assessed the impact of the presence of tau on connection strength at a variety of spatial scales. First, we averaged weighted degree across the whole brain, resulting in a single measure for each individual. As the overall number of connections in each individual’s graph was thresholded
at an identical network density, this measure represented the average strength of the strongest x% of connections. To assess global disease burden in Alzheimer’s disease we averaged 18F-AV-1451 BP ND across the whole brain. In PSP, we averaged 18F-AV-1451 BP ND across midbrain and basal ganglia, reflecting the more focal distribution of disease and our hypotheses about dissociated functional effects of cortical and subcortical tau (Passamonti et al., 2017).

Confirmatory tests demonstrated that the pattern of the moderation analyses below were unchanged if whole-brain average 18F-AV-1451 BP ND was used to assess tau burden in PSP. Both methods of assessing tau burden were independently assessed in the control group, with no significant effects demonstrated in either case.

In Alzheimer’s disease, we found a negative correlation between average connection strength and tau burden (Pearson’s r = -0.58, P = 0.015) (Fig. 2D). In PSP this relationship was reversed, with average connection strength increasing in line with tau burden (r = 0.65, P = 0.004). No relationship was observed in controls (r = -0.10, P = 0.75). Moderation analysis confirmed a dissociated relationship between tau and connection strength in Alzheimer’s disease and PSP at all network densities [Δr² = 0.29, F(1,30) = 19.0, P = 0.0001; after partialling out age (Supplementary Fig. 5A) Δr² = 0.28, F(1,30) = 17.9, P = 0.0002]. Therefore, while in Alzheimer’s disease the presence of tau pathology causes the strongest functional connections to weaken, in PSP, the presence of tau had the opposite effect, i.e. to strengthen functional connections.

Second, we assessed the distribution of change by repeating the correlation of disease burden against weighted degree at every individual node. We hypothesized that this effect would be greatest in those regions that display the strongest functional connectivity in the healthy brain, and which we have demonstrated to accrue most tau in Alzheimer’s disease. The gradient of this nodewise relationship reflects a measure of local change in weighted degree with disease burden (Fig. 2E). First, we examined whether the whole-brain average relationship could be replicated in these individual gradients, by performing sign tests. For Alzheimer’s disease, a negative relationship was confirmed (Z = -13.0, P < 0.0001); for PSP there was a positive relationship (Z = 14.8, P < 0.0001); while for controls no relationship was demonstrated using either whole brain (Z = -1.1, P = 0.27) or deep brain (Z = 1.1, P = 0.27) tau burden.

Next, we assessed whether the functional connectivity change at each node related to local tau burden, by correlating the gradient of the disease-related change in weighted degree with the disease-associated increase in 18F-AV-1451 binding potential at each node (i.e. correlating each column of Fig. 2F with the corresponding column of Fig. 2G). A negative correlation between these measures was demonstrated in Alzheimer’s disease (Pearson’s r = -0.30,
P < 0.0001, Spearman’s rho = −0.24, P < 0.0001). This relationship was absent in PSP (Pearson’s r = −0.07, P = 0.11, Spearman’s rho = −0.00, P = 0.98) and in controls (Pearson’s r = −0.01, P = 0.75, Spearman’s rho = −0.01, P = 0.80). As well as being present at all examined network densities, the negative correlation between disease-related change in functional connectivity strength and tau burden in Alzheimer’s disease was replicated in nodal connectivity strength, the equivalent unthresholded measure (r = −0.34, P < 0.0001, Supplementary Fig. 4B).

Finally, we assessed whether the functional connectivity change at each node related to the strength of its connections in the healthy control brain (i.e. correlating each patient panel of Fig. 2G with the control panel in Fig. 2H). As would be expected from the propensity of highly connected nodes to accrue tau, a negative relationship was demonstrated in Alzheimer’s disease (Pearson’s r = −0.23, P < 0.0001, Spearman’s rho = −0.24, P < 0.0001). Importantly, however, this relationship explained less variance than AV binding, with which we have demonstrated it to be correlated. In PSP, a positive relationship was demonstrated (Pearson’s r = 0.27, P < 0.0001, Spearman’s rho = 0.26, P < 0.0001).

In summary, nodes that are constitutionally more strongly connected to a larger volume of cortex are more likely to accrue tau pathology in Alzheimer’s disease but not PSP. This relationship is independent of the connectivity network to which a node belongs. Once present, the tau pathology appears to cause local functional connectivity strength to fall. By contrast, in PSP tau selectively accumulates in midbrain and deep nuclei, which has the consequence of increasing the strength of cortico-cortical functional connectivity, especially in those nodes that are constitutionally highly connected.

**Reorganization of cortical functional connectivity reflects cortical versus subcortical pathology**

We hypothesized that the presence of subcortical pathology in PSP might be causing an increase in weighted degree by necessitating an increase in the relative strength of short-range cortico-cortical connections as longer range connections are disrupted. This view is supported by the observation that a consequence of increasing tau burden in PSP is a marked reduction in participation coefficient (Fig. 4D). This relationship was not observed in Alzheimer’s disease, where increasing tau burden caused increasing participation. Moderation analysis confirmed this differential relationship at all network densities from 1% to 10% [Δr² = 0.21, F(1,30) = 7.6, P = 0.002; after partialling out age (Supplementary Fig. 5B) Δr² = 0.17, F(1,30) = 8.3, P = 0.007].

We further examined this hypothesis by examining the effect of tau on other measures of graph structure in each individual, dissociated from variation in degree by binarization after thresholding and normalization against 1000 random graphs with the same number of connections at each node.

The clustering coefficient quantifies how many of a node’s neighbours are neighbours of each other. All individuals in all groups displayed clustering to at least 2.8 x that of random graphs of the same density (Fig. 5D), consistent with a small-world connectivity distribution. While in Alzheimer’s disease a non-significant trend was observed towards reduced clustering with increasing tau burden, in line with previous reports (Stam et al., 2006; Sanz-Arigita et al., 2010), in PSP the opposite relationship were observed. Moderation analysis trended towards significance at the density of primary interest [Δr² = 0.08, F(1,30) = 3.8, P = 0.06; after partialling out age (Supplementary Fig. 5C) Δr² = 0.07, F(1,30) = 2.5, P = 0.09], and was significant from 7% to 10% density. The betweenness centrality of a node is a measure of the number of shortest paths between any other two nodes that pass through it. If the presence of subcortical pathology in PSP means that long-range information transfer must occur through a transcortical route, average betweenness centrality should increase. Conversely, predominantly cortical pathology in Alzheimer’s disease might increase reliance on cortico-subcortical connections, reducing average betweenness centrality. This prediction was verified (Supplementary Fig. 6A), with a differential relationship between tau burden and betweenness centrality confirmed by moderation analysis at all network densities from 1–10% [Δr² = 0.21, F(1,30) = 9.3, P = 0.005; after partialling out age (Supplementary Fig. 5D) Δr² = 0.20, F(1,30) = 8.9, P = 0.006].

Long-range information transfer by a cortico-cortical route implies an inefficient, indirect process, perhaps accounting for the cognitive slowing characteristic of neurodegenerative disorders with predominant subcortical pathology. We tested for this by examining closeness centrality, which is the inverse of the path length between a node and all other nodes in the graph. As would be predicted by this account, increasing disease burden resulted in a higher average path length (lower closeness centrality) in PSP but a lower average path length in Alzheimer’s disease (higher closeness centrality) at network densities from 2% to 10% [Fig. 6A, moderation Δr² = 0.20, F(1,30) = 9.1, P = 0.005; after partialling out age (Supplementary Fig. 5E) Δr² = 0.16, F(1,30) = 8.0, P = 0.008]. Averaged across the whole brain, this measure is equivalent to the global efficiency of the graph, and indeed the same statistical results were obtained from a moderation analysis of that metric; increasing disease burden resulted in reduced global efficiency in PSP.

While tau-mediated reorganization of connectivity in PSP results in slower and less efficient long-range information transfer, in Alzheimer’s disease one might suppose it to be a beneficial consequence of the loss of overall connection strength as those connections that remain are more globally efficient. However, we hypothesized this would be at the cost of local efficiency, which is a measure of the number of
strong connections between neighbouring nodes and the robustness of local networks to disruption. Indeed this was found to be the case at network densities from 4% to 10% [Fig. 7A, $\Delta r^2 = 0.15$, $F(1,30) = 6.5$, moderation $P = 0.016$; after partialling out age (Supplementary Fig. 5F) $\Delta r^2 = 0.12$, $F(1,30) = 5.5$, $P = 0.026$].

Finally, we assessed the functional influence of each node on the other nodes in the network by examining the eigenvector centrality. Again, an opposite effect of disease burden was observed in Alzheimer’s disease and PSP at network densities from 3–10% [Supplementary Fig. 7A, $\Delta r^2 = 0.22$, $F(1,30) = 8.5$, $P = 0.007$; after partialling out age (Supplementary Fig. 5G) $\Delta r^2 = 0.20$, $F(1,30) = 8.3$, $P = 0.007$] such that, on average, as PSP progressed each node had less functional influence on every other node in the graph.

Figure 4 Weighted participation coefficient. (A–C) Group-averaged weighted participation coefficient at each node, plotted against $^{18}$F-AV-1451 binding potential at that node. A negative relationship was observed in Alzheimer’s disease (AD), in violation of the metabolic demand hypothesis. (D) Between-subjects analysis of the relationship between global tau burden and participation coefficient at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (Alzheimer’s disease and PSP) was statistically significant. (E) The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups. (F) Average $^{18}$F-AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. (G) The local tau burden-related change in participation coefficient is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. (H) Average raw values for participation coefficient within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.
Local functional connectivity reorganization is related to tau in Alzheimer’s disease

To understand the differential reorganization of functional connectivity in Alzheimer’s disease and PSP, we replicated the approach taken for weighted degree; correlating the gradient of the disease-related change at every individual node with local change in °F-AV-1451 binding potential.

![Figure 5 Clustering coefficient](https://academic.oup.com/brain/article-abstract/141/2/550/4775021)

Figure 5 Clustering coefficient. (A–C) Group-averaged clustering coefficient at each node, plotted against °F-AV-1451 binding potential at that node. A positive relationship was observed in Alzheimer’s disease (AD), in violation of the trophic support hypothesis. (D) Between-subjects analysis of the relationship between global tau burden and clustering coefficient at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups (Alzheimer’s disease and PSP) was statistically significant. (E) The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups. (F) Average °F-AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. (G) The local tau burden-related change in clustering coefficient is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. (H) Average raw values for clustering coefficient within each group. The control mean for each metric is subtracted at every node. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.
Supplementary Figs 6B and 7B). In all cases, controls showed either no relationship or a weaker effect in the same direction as patients with Alzheimer’s disease.

The reorganization of graph metrics followed two distinct patterns. Closeness centrality (Fig. 6D) displayed a global effect, with most nodes increasing in Alzheimer’s disease and decreasing in PSP. This was not strongly related to local $^{18}$F-AV-1451 binding potential in Alzheimer’s disease (Pearson’s $r = 0.05$, $P = 0.25$, Spearman’s rho = 0.06, $P = 0.18$) or PSP (Pearson’s $r = 0.08$, $P = 0.06$, Spearman’s rho = 0.11, $P = 0.01$). Similarly, participation coefficient displayed global changes in both diseases (Fig. 4G) (Alzheimer’s disease Pearson’s $r = 0.09$, Spearman’s rho = 0.07, PSP Pearson’s $r = 0.08$, Spearman’s rho = 0.06).

By contrast, eigenvector centrality (Supplementary Fig. 7D) was more strongly related to local $^{18}$F-AV-1451 binding potential in Alzheimer’s disease (Pearson’s $r = -0.28$, $P < 0.0001$, Spearman’s rho = $-0.25$, $P < 0.0001$). Strikingly, the positive relationship we demonstrated across the whole brain masked opposing regional effects. As global tau burden increased, the functional influence of frontal regions on all other regions increased, while that of occipital regions decreased. In PSP, no consistent relationship with $^{18}$F-AV-1451 was observed (Pearson’s $r = -0.10$, $P = 0.02$, Spearman’s rho = $-0.04$, $P = 0.38$), with almost all regions having less functional influence on all other regions as tau burden increased.

Clustering coefficient (Fig. 5G), local efficiency (Fig. 7D) and betweenness centrality (Supplementary Fig. 6D) displayed an intermediate degree of regional specificity, being weakly but significantly correlated with change in $^{18}$F-AV-1451 binding potential in Alzheimer’s disease (clustering coefficient Pearson’s $r = -0.12$, $P < 0.0001$, Spearman’s rho = $-0.14$, $P < 0.0001$, local efficiency Pearson’s $r = -0.16$, $P < 0.0001$, Spearman’s rho = $-0.19$, $P < 0.0001$, betweenness centrality Pearson’s $r = -0.19$, $P < 0.0001$, Spearman’s rho = $0.19$, $P < 0.0001$) and PSP (clustering coefficient Pearson’s $r = -0.14$, $P < 0.0001$, Spearman’s rho = $-0.12$, $P < 0.0001$, local efficiency Pearson’s $r = -0.14$, $P = 0.001$, Spearman’s rho = $-0.11$).
Discussion

We have demonstrated that in Alzheimer’s disease a strong relationship exists between the propensity of a node to display elevated $^{18}$F-AV-1451 binding and the volume of cortex to which it is strongly connected (Fig. 2A). Further, we have demonstrated that this effect exists both within and between intrinsic connectivity networks (Fig. 3). This is consistent with the theory of trans-neuronal spread. The predictions of the competing hypotheses that highly active brain regions are vulnerable to tau accumulation due to a positive relationship with metabolic demand (Fig. 4A) or negative relationship with clustering due to a lack of trophic support (Fig. 5A) were not supported by the data.

In contrast, in PSP, we have demonstrated the opposite findings. Those brain regions that accrue most tau display weak connectivity (Fig. 2B), but are predicted to have high metabolic demand (Fig. 4B) and a lack of trophic support (Fig. 5B).

Further, we have explored the consequences of tau accumulation with cross-sectional analyses at a variety of spatial scales. In Alzheimer’s disease, we have demonstrated that with greater levels of tau pathology the strongest internodal connections are weakened (Fig. 2G). This reorganization of the brain network leads to more direct long-range connections passing through fewer nodes (Fig. 6A and Supplementary Fig. 6A), at the cost of lower local efficiency (Fig. 7A).

In PSP, where tau accumulation is predominantly subcortical, we demonstrate the opposite reorganization of the connectivity graph. With greater levels of tau in midbrain and deep nuclei (Fig. 2F), the strongest functional connections are strengthened; these are predominantly cortical (Fig. 2G and H). Information transfer therefore takes a less direct path (Supplementary Fig. 6A), passing through a larger number of cortical nodes en route (Fig. 4A), as deep structures can no longer sustain long range $P = 0.02$, betweenness centrality Pearson’s $r = 0.17$, $P < 0.0001$, Spearman’s rho = 0.08, $P = 0.08$).

Figure 7 Local efficiency. (A) Between-subjects analysis of the relationship between global tau burden and local efficiency at a network density of 6%. No significant relationship was found for the control subjects with either method of assessing tau burden. Moderation analysis for a differential relationship between graph metric and tau burden in the two disease groups [Alzheimer’s disease (AD) and PSP] was statistically significant. (B) The magnitude of disease related change at each node is plotted as a single point, grouped by lobe. Control data are shown for both methods of assessing tau burden. Stars represent statistically significant excesses of positive or negative gradients in the disease groups (Supplementary Table 2). (C) Average $^{18}$F-AV-1451 binding potential at each node. Red spheres represent increases compared to the cerebellar reference region. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the magnitude of binding at that node. All groups are scaled identically. (D) The local tau burden-related change in local efficiency is plotted for each group at each node. Red spheres represent local increases as a result of greater overall tau burden; blue spheres represent local decreases. The radius of each sphere is linearly related to the magnitude of disease-related change at that node across the whole range of disease burden observed in each group. All groups are scaled identically. Control images are based on the whole-brain method of assessing tau burden. (E) Average raw values for local efficiency within each group. Red spheres represent increases compared to the control mean value. Blue spheres represent decreases. The centre of the sphere is placed at the centre of the region of interest. The radius of each sphere is linearly related to the difference from the control mean at that node. All groups are scaled identically.
connectivity. This is accompanied by a decrease in participation coefficient (Fig. 4D) and an increase in clustering (Fig. 5D) as connectivity becomes increasingly modular.

**Insights into the mechanisms of disease progression in humans**

It has been proposed that the pathological mechanisms underling Alzheimer’s disease begin in a single, vulnerable location and spread from cell to cell, rather than occurring independently in a large number of vulnerable cell populations (Guo and Lee, 2014; Goedert, 2015). The primary direct evidence for such propagation of tau comes from rodent studies. For example, the injection of brain extract from transgenic mice expressing mutant tau into mice expressing wild-type human tau caused wild-type tau to form filaments and spread to neighbouring brain regions (Clavaguera et al., 2009). Further, pathological tau from human brains causes disease in wild-type mice, in which the pathological human tau species becomes self-propagating (Clavaguera et al., 2013). This tau propagation is mediated by the presence and strength of synaptic connectivity rather than spatial proximity (Liu et al., 2012; Iba et al., 2013; Ahmed et al., 2014).

Associative studies of the healthy brain have demonstrated that large-scale, functionally connected neural networks strongly resemble the known patterns of atrophy in distinct neurodegenerative syndromes mediated by tau and TDP-43 (Zhou et al., 2012). Further human evidence comes from the observation that patterns of atrophy in the rare disease non-fluent variant primary progressive aphasia (nfvPPA) strongly correspond to structural and functional connectivity in the healthy speech production network (Mandelli et al., 2016). This is an important observation, because clinically and radiologically indistinguishable cases of nfvPPA can be caused either by tau or by the unrelated protein TAR DNA-binding protein 43 (TDP-43), which has been demonstrated to propagate trans-neuronally (Braak et al., 2013).

Here, we go beyond these associative studies to measure tau burden and functional connectivity in the same individuals at both the whole-brain and regional level. Our observation that those brain areas that are more strongly functionally connected have accrued more tau pathology in Alzheimer’s disease (Fig. 2A), independent of which connectivity network they belong to (Fig. 3), is consistent with trans-neuronal spread. We demonstrate that the presence of tau is not, in itself, inducing stronger regional connectivity by our cross-sectional analysis of the Alzheimer’s disease group, in which we demonstrate that as cortical tau accumulates the overall functional connectivity of cortex falls (Fig. 2D), and this between-subjects effect is strongest in those brain regions with most tau accumulation (Fig. 2F and G). Crucially, we demonstrate that 18F-AV-1451 binding potential at each node is better than the connectivity of that node in the healthy brain at accounting for regional variance in connectivity change, arguing against the presence of tau being a secondary marker of neurodegeneration in vulnerable hubs. In other words, it is not coincidence that Alzheimer’s disease tends to impact large networks; it is a predictable consequence of trans-neuronal spread of a disease-causing protein. Graph theoretic models of transmissible disease epidemics are in agreement that the likelihood of an individual becoming infected (and the dose of the infectious agent received) is directly proportional to its number of infected neighbours and their infectivity (Durrett, 2010). As our nodes represent brain regions of equal volume, the binary portion of degree represents a surrogate measure of the number of neurons to which a brain region is connected, and the weighted portion of degree is a measure of the strength of these connections. By the time Alzheimer’s disease is sufficiently advanced to cause the symptoms of MCI, tau is generally already present to some degree throughout the neocortex (Markesbery, 2010), and therefore reaching a disease stage at which the number of neighbours more closely approximates the number of infected neighbours, and the connection strength between infected neighbours (here the weighted portion of degree) becomes a strong driver of infectivity.

Conversely, our analyses do not provide support for models of hub vulnerability due to metabolic demand or lack of trophic support in Alzheimer’s disease. It is important to acknowledge that this does not mean that these mechanisms are unimportant, but rather that they are a downstream event of tau accumulation. In other words, while we demonstrate that the propensity of a node to accrue tau is not related to metabolic demand or trophic support, these factors might still contribute to determining the vulnerability of brain regions to the presence of a given amount of tau. This hypothesis could be addressed in future studies by relating the information content of tau ligand binding to other measures of neurodegeneration such as longitudinal changes in grey matter volume.

In PSP, we do find support for models of hub vulnerability due to metabolic demand and lack of trophic support, but not for models of trans-neuronal spread. This is in line with other recent studies, demonstrating that neurodegeneration in PSP is related to local gene expression patterns (Rittman et al., 2016).

**Is the difference between Alzheimer’s disease and PSP mediated by tau isoform or intrinsic connectivity?**

A striking feature of our results is that we demonstrate a strong relationship between 18F-AV-1451 binding potential and the strength of functional connectivity in Alzheimer’s disease but not PSP. There are at least two potential (and not mutually exclusive) explanations for this dissociation.

The first possibility is that trans-neuronal tau propagation might occur more slowly in PSP. Cellular models have demonstrated that the propensity of tau to propagate...
Intracellularly depends on its ability to form aggregates (Falcon et al., 2015). Once present in a new cell, the ability of tau fibrils to induce the aggregation of constitutionally present tau depends on the conformation of the fibril structure (Nonaka et al., 2010; Fitzpatrick et al., 2017). It might be that the straight filaments of predominantly 4R tau that characterize PSP are simply less able to propagate or need to be present in higher concentrations before they can induce a chain reaction of local tau aggregation (Guo et al., 2016). This view is supported by our demonstration that tau is restricted to brainstem and deep nuclei in PSP despite these nodes being highly promiscuous between networks, displaying a high participation coefficient (Fig. 4B) and low clustering (Fig. 5B).

Second, it is possible that tau propagation does occur, but that it is limited in range to a subcortical intrinsic connectivity network (Raj et al., 2012) (perhaps network 5 in Laird et al., 2011). Such a subcortical network may be poorly visualized by multi-echo functional MRI, and is less frequently observed at rest than the DMN (Greicius et al., 2003), which has been implicated in Alzheimer’s disease (Greicius et al., 2004). Together, therefore, our findings of weak functional connectivity in deep nuclei and brainstem nodes (Fig. 2B) might mask meaningful functional connectivity within and between these regions, accounting for the restricted and stereotyped pattern of tau accumulation in early PSP (Ellison et al., 2012) and cortical escape in advanced PSP (Schofield et al., 2012).

Cross-sectional data reveals patterns less visible at the group level

Computational modelling of connectivity-dependent cell death predicts a dissociation between changes in functional connectivity in early and late Alzheimer’s disease (de Haan et al., 2012). It has been proposed that, in early disease, hubs compensate for declines in structural connectivity by increasing their firing rate, manifesting as stronger functional connectivity (Maestú et al., 2015). As disease progresses, this mechanism breaks down as neural damage prevents the maintenance of this metabolically demanding compensation (Jones et al., 2015). This dissociation is thought to underlie some of the seemingly inconsistent findings in the analysis of graph properties in neurodegenerative disease. By using a cross-sectional approach across a range of disease severity, we demonstrate relationships consistent with this hypothesis. In Alzheimer’s disease, weighted degree (Fig. 2D) and betweenness centrality (Supplementary Fig. 6A) consistently fall as tau burden increases. However, in early disease where tau burden is low (our sample includes a range of severity, including PB-positive MCI), the regression line for these metrics is above the control average. Similarly, examining the regional changes related to disease burden reveals striking patterns that are obscured at the group level. This underlines the more general principal that one should be cautious in interpreting a main effect of group in the presence of an interaction, or correlation with severity.

Reorganization of brain networks

Our examination of two distinct tau-mediated neurodegenerative pathologies with different distributions of pathology has enabled us to distinguish their consequences. We demonstrate strongly opposing effects in a range of metrics resulting from the presence of predominantly cortical (Alzheimer’s disease) or subcortical (PSP) tau. The direction of these effects in Alzheimer’s disease is consistent with the previously recognized impact of neurodegeneration, leading to increasingly random cortical connectivity (Sanz-Arigita et al., 2010; Stam, 2014) and a reduction in small-world properties (Stam et al., 2006). Eigenvector centrality showed particularly strong regional effects in Alzheimer’s disease, with a negative correlation observed between disease-related changes in this metric and local tau burden at each node—i.e. those brain regions that displayed less tau pathology had greater functional influence on other brain regions.

In contrast, cortico-subcortical functional connectivity is preferentially impaired in PSP-Richardson’s syndrome, resulting in cerebral information transfer taking a less direct path through a larger number of cortical nodes, reducing closeness centrality and eigenvector centrality, but increasing cortical degree, betweenness centrality and local efficiency. This results in an excessively modular connectivity arrangement, with decreasing participation coefficient and increasing clustering. These findings tie together classical observations of the ‘subcortical dementia’ phenotype of PSP (Albert et al., 1974) with more modern observations that behavioural change and cognitive impairment in PSP correlates with frontal cortical hypometabolism (D’Antona et al., 1985; Foster et al., 1988) and atrophy (Cordato et al., 2002, 2005). We propose that increases in cortical functional connectivity can compensate for subcortical tau burden in PSP. Increasingly indirect information transfer accounts for the cognitive slowing that is the hallmark of the ‘subcortical dementias’, but performance on untimed tests is preserved until cortical regions become atrophic in late disease.

Study limitations

The main limitation of our analysis is that it is cross-sectional, and we use 18F-AV-1451 binding as a surrogate marker of tau burden. By making observations about the relationship between tau burden and functional connectivity in this way, we assume a uniformity of effect within our disease groups. As novel PET ligands such as 18F-AV-1451 gain maturity, longitudinal assessment of tau burden and functional connectivity in the same individuals will be an important and powerful validation of our results. Definite evidence of the causal relationship between tau and connectivity with require the combination of longitudinal assessment and interventional studies targeting tau pathology.
It should also be noted that $^{18}$F-AV-1451 binding identifies predominantly aggregated tau in tangles, and does not directly measure oligomeric tau, which may be more toxic to the cell and synaptic plasticity, nor extracellular forms of tau that may mediate spread of pathology. The molecular binding target of $^{18}$F-AV-1451 in non-Alzheimer’s disease tauopathies is disputed; elevated ‘off-target’ binding has been demonstrated in the basal ganglia of healthy controls (Johnson et al., 2016), albeit to a lesser degree than that observed in PSP (Fig. 2F) (Passamonti et al., 2017), and in TDP-43 associated disorders without evident tau pathology (Bevan-Jones et al., 2017a). Nonetheless, $^{18}$F-AV-1451 is able to recapitulate the distribution of post-mortem neuropathology in these disorders, making it appropriate for use here (Smith et al., 2016).

Our analysis is focused towards cortico-cortical functional connectivity. In particular, multi-echo MRI might have a poor signal-to-noise ratio in deep brain structures, although the main advantage of using this sequence is that it enables robust de-noising of movement-related artefacts pipeline (Kundu et al., 2012, 2013). This is critical in clinical populations, in which functional MRI data may differentially suffer from quality degradation due to head movements.

Finally, by examining proportionately thresholded graphs with 1–10% density, our analysis focuses on the strongest interregional functional connections. However, it is possible that we are missing additional effects of neurodegeneration on weak or medium-strength connections. Tract-tracing studies indicate that there are weak anatomical connections, equivalent to a few axons, between some cortical areas (Ypma and Bullmore, 2016). Such weak links may have functional importance in complex networks (Granovetter, 1983). However, weak connections are difficult to evaluate with functional MRI, as it is not possible to disentangle them from correlation arising from signal noise. Future evaluation of these weaker connections with in vivo tractography, neuropathology or novel methods might reveal additional effects not evident in our dataset. In the interim, the thresholding procedure retains several advantages; by retaining only the most strongly correlated edges one is less likely to include false positive correlations and topologically random edges. It also allows the computationally intense process of normalization of metrics against random graphs of equal density. The consistency between the results using thresholded nodal weighted degree and unthresholded nodal connectivity strength provides reassurance in the choice of thresholding of connections.

**Conclusion**

This study reveals the differential relationship between tau burden and functional connectivity in two distinct human neurodegenerative tauopathies. Our results enable us to disentangle the causes of tau accumulation from their consequences. They have wide-ranging implications, from the validation of models of tau trafficking in humans to corroborating computational models of hub compensation in Alzheimer’s disease, while accounting for the contrasting cognitive phenotype of these two conditions. These insights into the relationship between tau burden and brain connectivity changes will inform translational models and clinical trials of disease-modifying therapies.

**Acknowledgements**

We would like to thank our volunteers for participating in this study, the radiographers/technologists at the Wolfson Brain Imaging Centre and PET/CT Unit, Addenbrooke’s Hospital, for their invaluable support in data acquisition, Young T. Hong and Tim D. Fryer for their assistance in modelling the PET data, Franklin I Aigbirhio for radiogand production, and Edward T. Bullmore for advice on interpretation and control analyses. We thank Avid (Lilly) for supplying the precursor for the manufacturing of $^{18}$F-AV-1451 for use in this study.

**Funding**

The NIMROD study was funded by the National Institute for Health Research (NIHR, RG64473) Cambridge Biomedical Research Centre and Biomedical Research Unit in Dementia, PSP Association, the Wellcome Trust (JBR 103838), the Medical Research Council (MC-A060-5PQ30). T.E.C. is supported by a personal fellowship from the Association of British Neurologists and Patrick Berthoud charitable trust.

**Supplementary material**

Supplementary material is available at *Brain* online.

**References**


Evidence for causal top-down frontal contributions to predictive processes in speech perception

Thomas E. Cope¹, E. Sohoglu², W. Sedley³, K. Patterson¹⁴, P.S. Jones¹, J. Wiggins¹, C. Dawson¹, M. Grube³, R.P. Carlyon², T.D. Griffiths³, Matthew H. Davis² & James B. Rowe¹,²

Perception relies on the integration of sensory information and prior expectations. Here we show that selective neurodegeneration of human frontal speech regions results in delayed reconciliation of predictions in temporal cortex. These temporal regions were not atrophic, displayed normal evoked magnetic and electrical power, and preserved neural sensitivity to manipulations of sensory detail. Frontal neurodegeneration does not prevent the perceptual effects of contextual information; instead, prior expectations are applied inflexibly. The precision of predictions correlates with beta power, in line with theoretical models of the neural instantiation of predictive coding. Fronto-temporal interactions are enhanced while participants reconcile prior predictions with degraded sensory signals. Excessively precise predictions can explain several challenging phenomena in frontal aphasias, including agrammatism and subjective difficulties with speech perception. This work demonstrates that higher-level frontal mechanisms for cognitive and behavioural flexibility make a causal functional contribution to the hierarchical generative models underlying speech perception.
It has long been recognised that perception relies on the integration of sensory input with expectations based on prior knowledge or experience. This can be instantiated in hierarchical generative models, which contain both top-down connections for priors or beliefs about sensory evidence, and bottom-up connections for prediction error. The layers of these hierarchical models represent progressively more abstract descriptions of the underlying sensory data. An influential implementation is known as predictive coding, in which the top-down generative connections express predictions for expected sensory signals, while bottom-up processes pass forward prediction errors to update the model. This method of information transfer is highly efficient. Neural models of predictive coding are well formalised, and we therefore conceptualise and interpret our study in this framework.

There is empirical evidence for predictive coding in health, for vision, hearing, and the link between perception and action in motor control. Furthermore, dysfunctional predictive coding mechanisms can explain a range of neurological and psychiatric phenomena, in schizophrenia, functional movement disorders, alien limb syndrome, tinnitus and hallucinations. Although these disorders have been explained in terms of aberrant predictive coding, the functional consequences of degradation of the neural architecture responsible for generating top-down predictions are unknown. This is a critical and novel test for hierarchical models of perception, which motivates the following hypothesis: degeneration of top-down prediction mechanisms in frontal lobe should have a substantial impact on lower-level sensory responses in temporal lobe, and should impair perceptual function when prior knowledge and sensory input must be combined (Fig. 1).

We test this hypothesis in the context of speech perception. Speech is a natural domain in which to study prediction, as humans are able to exploit a wide variety of visual, contextual and semantic cues to improve perception, especially in difficult listening environments. Indeed, contradictory beliefs established by mismatching visual and auditory speech can lead to false perception. It is important to note that such multimodal integration can be modelled in terms of predictive coding regardless of whether or not visual information occurs before auditory information; what is important is that auditory sensory predictions are set up based on information from prior experience, sentential context or sensory information from another domain. We exploited the importance of written text in supporting perception of degraded speech. There is evidence for left lateralised top-down information transfer from frontal language and motor speech regions to auditory cortex during speech perception; in predictive coding theory this top-down transfer generates prior expectations for speech content and explains how listeners combine prior knowledge and sensory signals during perception and perceptual learning.

To assess the effects of disrupted predictions we studied patients with early non-fluent primary progressive aphasia (nvfPPA), which is associated with selective neurodegeneration of the frontal lobe language and motor speech areas, but preservation of temporal lobe auditory regions. Disordered speech output in nvfPPA is characterised by apraxia of speech and/or agrammatism. In contrast to stroke aphasia the neural damage

---

**Fig. 1** An illustration of the experimental motivation. **a** A schematic Bayesian framework for predictive coding in speech perception. **b** The putative brain basis of this framework. Predictions are generated in inferior frontal gyrus and/or frontal motor speech regions (pink), and instantiated in auditory regions of superior temporal lobe (pale blue). **c** The two dimensional experimental manipulation employed here to detect a dissociation between normal temporal lobe responses to sensory detail (number of vocoder channels) and abnormal frontal lobe responses to prior congruency. **d** Our experiment relies on detecting the consequences of degraded predictions in abnormal frontal brain regions by measuring their effects in normal temporal regions. **e** Voxel-based morphometry in our patient group. Regions coloured in red displayed consistent reductions in grey matter volume (FWE $p < 0.05$). Regions coloured blue had strong evidence for normal cortical volume in nvfPPA (Bayesian probability of the null $>0.7$, cluster volume $>1$ cm$^3$). Uncoloured (grey) areas had no strong evidence for or against atrophy.
in frontal regions is partial\(^36, 37\), enabling us to study a disruption of predictive mechanisms, rather than a system reorganised following their complete absence. Additionally, this patient cohort presents fewer problems for the modelling and interpretation of magnetoencephalography (MEG) or electroencephalography (EEG), as atrophy is subtle in early nfvPPA\(^{38, 39}\).

For patients and matched control participants we recorded behavioural and neural data showing the influence of top-down and bottom-up manipulations on speech perception using an established paradigm involving presentation of written text that matches or mismatches with degraded spoken words\(^{27, 32, 40}\). With this paradigm, we can determine whether and how fronto-cortical neurodegeneration impairs speech perception. The presence and function of top-down influences on speech perception is controversial (see refs. \(^{41–44}\)), as is the question of whether fronto-cortical regions make a critical contribution to speech perception, through predictive coding or alternative mechanisms. Some authors suggest that these contributions are task-specific and not a core component of speech perception systems\(^{35}\). The present study provides causal neural evidence with which to assess both of these claims.

Here we demonstrate distant neural effects of the degeneration of top-down signals from frontal lobes, during speech perception. We provide evidence of a direct relationship between the degree of frontal lobe degeneration and a delay in the neural mechanism for the reconciliation of predictions, which results in their inflexible application. Bayesian perceptual inference simulations demonstrate that this results in aberrantly precise prior expectations, which manifest as increased beta power during the instantiations of predictions, in agreement with theoretical frameworks of predictive coding. Finally, we show task-dependent enhancements in fronto-temporal interaction in nfvPPA, reflecting degraded neural mechanisms working harder to reconcile excessively precise predictions. We explain how inflexible predictions are able to account for several previously poorly understood symptoms and signs in frontal non-fluent aphasias, including difficulties with parsing the structure and content of running speech. Together, our results provide causal evidence for a critical role of frontal regions for the reconciliation of predictions during the perception and comprehension of speech.

### Results

#### Structural consequences of nfvPPA

To confirm the dissociation between frontal atrophy and intact temporal cortex, upon which our experiment relies, we used voxel-based morphometry to compare grey matter volume in nine of our patients with nfvPPA (see Table 1 for participant characteristics) to 36 age-matched healthy individuals using whole brain statistical parametric mapping (SPM) t-test and Bayesian null tests. As anticipated, brain regions of interest displayed a localised pattern of atrophy in nfvPPA (Fig. 1e), with grey matter volume loss in left inferior frontal regions (family wise error corrected (FWE) peak \(p = 0.001\) at montreal neurological institute (MNI) \([-37, 17, 7]\); Bayes posterior probability of no difference \(<0.00001\)), but not in left primary auditory cortex (FWE \(p = 1\); Bayes posterior probability of no difference 0.75 at MNI \([-59, -24, 9]\) or superior temporal gyrus (FWE \(p = 1\); Bayes posterior probability of no difference 0.91 at MNI \([-67, -17, 3]\)). Significant atrophy was also observed in right inferior frontal regions (FWE \(p = 0.004\); peak MNI \([37, 20, 6]\)) but not right primary auditory cortex (FWE \(p = 1\) at MNI \([59, -24, 9]\) or superior temporal gyrus (FWE \(p = 1\) at MNI \([67 -17 3]\)). Significant atrophy in left inferior frontal regions lay within pars triangularis, pars opercularis and anterior insula in the Desikan–Killiany Atlas (Supplementary Table 1).

### Subjective speech perception symptoms in nfvPPA

While the core symptoms in nfvPPA relate to apraxia of speech and agrammatism, patients often complain of a feeling of speech deafness. To test for this symptom in our cohort, we asked patients and controls to rate their subjective difficulty with five listening scenarios, by placing a mark on a line from ‘very easy’ to ‘very difficult’. Patients could respond appropriately to such rating scales. Patients and controls displayed very similar subjective difficulty ratings for ‘speech in noise’, ‘localising sounds’, ‘understanding station announcements’ and ‘how loud others say their television is’ (all \(t(20) > 0.3\), Supplementary Fig. 1A). However, there was a difference in their assessments of difficulty in understanding speech in quiet environments (\(t(20) = 2.66, p = 0.015\)); controls universally rated this as very easy, while patients rated it to be almost as difficult as understanding speech in noise (interaction \(F(1,20) = 8.21, p = 0.010\)). Patients and controls had similar, age-appropriate, hearing acuity (Supplementary Fig. 1B).

### Evoked neural responses during the reconciliation of predictions

To assess the neural correlates of degraded predictive mechanisms in nfvPPA (Fig. 1d), we recorded simultaneous MEG and EEG during a speech perception task. We manipulated prior expectations using matching or mismatching text cues, before participants heard spoken words that were varied in sensory detail by manipulating the number of vocoder channels (Figs. 1c, 2a)\(^{46}\). Overall evoked power was similar for the two groups of participants (Supplementary Fig. 3); fronto-neurodegeneration did not lead to any large difference in the magnitude of the neural response evoked by single spoken words that could manifest as spurious group by condition interactions in neural activity.

To confirm that patients have normal responses to manipulations of sensory detail independent of predictions (as expected given preserved cortical volume in auditory cortex and superior temporal gyrus, Fig. 1e), we first assessed the neural effect of the number of vocoder channels (Fig. 3). Across all 21 individuals, SPM F-test peak effects were observed in the planar gradiometers at 96 ms (scalp-time FWE \(p < 0.001\), in magnetometers at 188 ms (FWE \(p < 0.001\)) and again at 380 ms (FWE \(p < 0.001\)), and in EEG at 392 ms (FWE \(p < 0.001\)). These findings are consistent with previous studies in young individuals\(^{40}\). No reliable group by sensory interactions were found at the scalp locations of the peak main effect or at all scalp-time locations. Together, these results

---

### Table 1 Demographic details of the experimental groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Age</th>
<th>Gender</th>
<th>Age leaving education</th>
<th>MMSE</th>
<th>ACE-R</th>
<th>Raven’s matrices</th>
</tr>
</thead>
<tbody>
<tr>
<td>nfvPPA</td>
<td>11</td>
<td>72</td>
<td>9</td>
<td>18 (3)</td>
<td>28</td>
<td>84</td>
<td>36</td>
</tr>
<tr>
<td>MEG controls</td>
<td>11</td>
<td>72</td>
<td>9</td>
<td>17 (2)</td>
<td>29</td>
<td>95</td>
<td>46</td>
</tr>
<tr>
<td>MRI controls</td>
<td>36</td>
<td>73</td>
<td>9</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Mean (standard deviation). There were no statically significant differences in age, gender or education between nfvPPA patients and MEG controls. nfvPPA patients scored more poorly than controls on the Addenbrooke’s Cognitive Examination (Revised) and Raven’s Progressive Matrices, but were still within the population normal range. Most of the difference between nfvPPA and controls on the ACE-R was accounted for by verbal fluency. Audiometric thresholds are available in Supplementary Fig. 1B. One patient was unable to tolerate the MEG scanner environment so, for that case, results contribute only to the behavioural analysis.
are consistent with the idea that patients and controls produce similar neural responses to manipulations of sensory detail in degraded speech. Crucially for the interpretation of later results, the latency of these responses was also the same in both groups. In magnetometers, where the late effects of vocoder channel number were most clearly seen, control peaks occurred at 196 ms and 340 ms, and patient peaks at 172 ms and 372 ms (Fig. 3).

Having performed these control analyses, we tested our primary hypothesis by examining the neural effect of manipulating whether prior expectations matched sensory input. Across all 21 participants, during the reconciliation of predictions there were significant SPM F-test main effects of cue congruency in all sensor types (Fig. 4a); planar gradiometers at 464 ms (scalp-time peak FWE $p < 0.001$), magnetometers at 400 ms (FWE $p < 0.001$), and EEG at 700 ms (FWE $p < 0.001$). At these scalp locations, significant group by congruency interactions were observed in the planar gradiometers and in the magnetometers (unpaired t(19), $p < 0.05$ sustained over at least eight sequential samples), but not in the EEG electrodes. In the planar gradiometers, between 264 ms and 464 ms controls had a significantly larger effect of congruency than patients (Fig. 4b). The scalp topography averaged across this time window resembled that observed during the peak of the main effect, but with the pattern being stronger in controls (Fig. 4c). In the magnetometers, a cluster with similar timing and scalp topography was observed between 240 and 560 ms (Fig. 4b). Two additional clusters were also observed. In later time windows, from 728 to 808 ms, group by congruency interactions were observed in the opposite direction, with patients showing a significantly greater effect than controls. Again, the scalp topographies in this cluster resembled those during the main effect (Fig. 4e). This indicates that the effect of congruency was present in both groups, but that the effect was significantly delayed in nfvPPA. Finally, an early cluster was observed between 152 ms and 224 ms, with the controls displaying a significantly greater effect of congruency than patients. Intriguingly, the scalp topography in this time window was different to that observed during the conjoint main effect, with dipoles having a much more anterior centre of mass (Fig. 4c). This anterior topography is consistent with a frontal source, expected to appear in this earlier time window as shown in similar previous studies with young healthy listeners.

To assess the underlying neural sources of these effects, multimodal sensor data were combined and inverted into source space with sLORETA. For the main effect of vocoder channel number, reconstructions were performed across all individuals combined, because no group difference or group by clarity interaction was demonstrated in sensor space. The main effect of sensory detail shown in MEG sensors and EEG electrodes is explained by increased activity for 16 channel effects of sensory detail shown in mid-latency time windows (200–280 ms and 290–440 ms; Fig. 5a), replicating previous findings in younger individuals.

To localise the group by cue congruency interaction, we first display source reconstructions for the main effect of cue congruency in each group (Fig. 5b) for time windows defined by the main data features in overall sensor power averaged over conditions and participants (Supplementary Fig. 3). Given our findings of delayed congruency effects in patients, an additional, late, time window (710–850 ms) was also examined post hoc. We focus on the two principal sources observed in young healthy individuals for this task, extracting average power in each
**Fig. 3** The effect of vocoder channel number. Illustrative topographic plots are shown of the main effect of vocoder channels across all participants. No group by sensory detail interactions were observed either at the peak locations (marked by white stars) or in a confirmatory SPM analysis.

**Fig. 4** The effect of prime congruency. **a** Illustrative scalp topographic plots of the main effect of cue congruency for each group from 400 ms to 700 ms, a period of where both groups showed a large statistical effect of congruency with similar topography. White stars indicate the scalp location of the peak congruency effect across both groups between ~100 ms and 900 ms (FWE $p < 0.001$ for all sensor types). **b** Significant group by congruency interactions ($p < 0.05$ sustained for more than 25 ms at the scalp locations marked by white stars in the upper panel) were observed in planar gradiometers and magnetometers, and are shaded in lilac. **c** Topographic plots for each group are shown averaged across each significant cluster of group by congruency interaction.
condition at frontal and superior temporal voxels of interest defined by the main effect of congruency averaged across the whole epoch (Fig. 5d). No group differences were demonstrated in the earliest (90–150 ms frontal $F(1,18) = 0.06, p = 0.81$, temporal $F(1,18) < 0.01, p = 0.99$) or later (450–700 ms frontal $F(1,18) = 4.24, p = 0.054$, temporal $F(1,18) = 2.84, p = 0.11$, 710–860 ms frontal $F(1,18) = 3.10, p = 0.10$, temporal $F(1,18) = 2.96, p = 0.10$) time windows. Between 200 and 280 ms, there were significant main effects of group in both frontal ($F(1,18) = 6.37, p = 0.02$) and temporal ($F(1,18) = 5.07, p = 0.04$) voxels, with greater responses in controls than patients. Between 290 and 440 ms, this main effect had dissipated (frontal $F(1,18) = 2.51, p = 0.13$, temporal $F(1,18) = 1.63, p = 0.21$), but there was a group by condition interaction, with controls showing a greater effect of cue congruency in the superior temporal ($F(1,18) = 4.46, p = 0.049$) but not the frontal ($F(1,18) = 0.76, p = 0.39$) voxel.

The analysis across the whole epoch is of particular interest. Across all individuals, a repeated-measures ANOVA (Supplementary Table 2) confirmed the pattern of opposing effects of prior knowledge in frontal and superior temporal regions seen in a previous study $^{40}$ ($F(1,134) = 60.1, p < 0.001$). However, there was a group by source by congruency interaction ($F(1,134) = 11.9, p = 0.001$), primarily driven by the absence of a significant effect of congruency in frontal regions in nfvPPA (Fig. 5b). When the total power in the frontal region was examined, a main effect of group was observed such that patients had significantly more frontal power than controls, but their modulation of frontal power by congruency was absent (Fig. 5d).

**Behavioural experiment 1 vocoded word clarity rating.** Given these neural differences, we sought to understand the perceptual correlates of neural delay in the reconciliation of predictions by examining the behavioural consequences of manipulations of prior knowledge in our two groups (Figs. 2a, 1c). All individuals reported that the perceptual clarity of vocoded words was significantly increased by matching text cues (Fig. 2b), but this effect was greater in patients with nfvPPA than in controls. A repeated-measures ANOVA revealed that Group, Number of Vocoder Channels and Cue Congruency were significant either as main effects or as part of two-way or three-way interactions (see Table 2, Experiment 1 for statistical details).

A replication experiment outside the MEG scanner confirmed that the difference between match and mismatch trials was due to a facilitatory effect of matching prior knowledge and not simply increased confusion in the face of mismatching priors: ratings of
perceptual clarity after a mismatching text cue were not statistically different from those after a ‘neutral’ or uninformative cue (repeated measures ANOVA $F(1,120) = 2.09, p = 0.15$). Furthermore, patients with nfvPPA had a much larger difference in clarity rating between ‘neutral’ and ‘match’ trials compared to controls (Supplementary Fig. 1C). It is important to note that participants were explicitly instructed to rate clarity across their own range of perceptual experience within the experiment, and were given training until they were able to do this. Comparing clarity ratings across groups is not, therefore, a direct measure of comparative listening difficulty as a rating of ‘1’ simply means ‘one of the least clear words I heard in the experiment’, while a ‘4’ means ‘one of the clearest words I heard’. To fully assess the perceptual basis of our findings, we assessed the elements contributing to perceptual clarity with a further experiment and Bayesian modelling. These elements included: (1) patients’ and controls’ ability to identify degraded spoken words and (2) participants’ introspective ability to perform higher-level estimation of the global precision of sensory input.

**Behavioural experiment 2 vocoded word identification.** To ensure that our finding was not a consequence of impaired word identification in patients leading to a group difference in reliance on prior knowledge$^{36}$, we performed a second experiment in which participants identified noises vocoded words in the absence of prior expectations (Fig. 2c). To reduce response demands for patients with non-fluent speech we used a four-alternative forced-choice identification task. All individuals with nfvPPA were above chance at identifying even the most degraded vocoded speech and, as a group, performed almost as well as controls (Fig. 2d). Both groups were influenced in the same way by the number of noise vocoder channels and the number of close distractor items presented as alternatives in the forced choice. As expected, it was easier for all individuals to identify words with more vocoder channels and if there were fewer close distractor items. This effect was strongest for the most degraded speech, manifesting as an interaction between vocoder channels and distractor difficulty (Table 2, Experiment 2).

Crucially, these data show that a lower-level impairment in perceiving vocoded speech cannot be the sole explanation of our finding of an increased congruency effect in nfvPPA patients. Patients performed better at identifying speech with eight channels than controls did with four channels (repeated measures ANOVA $F(1,63) = 7.1, p = 0.015$). Yet, patients still display a larger congruency effect for 8-channel vocoded words than controls do for 4-channel speech ($t(20) = 2.17, p = 0.04$). Hence, the magnitude of congruency effects in clarity rating is not simply related to objective abilities at word identification, but rather reflects a difference in the mechanisms by which prior knowledge influences lower-level perceptual processing. We investigate the nature of this effect with a Bayesian perceptual model combining word report and clarity rating data.

**Bayesian modelling of experiments 1 and 2.** To dissociate changes in the precision of predictions from difficulties with higher level estimation of the precision of sensory input, we performed hierarchical Bayesian inference simulations (c.f. ref. $^{33}$; Supplementary Fig. 2). Individual differences in word discriminability were accounted for by defining the precision of sensory input for each subject as the percentage above chance for word identification at each vocoder channel number in Experiment 2. This allowed us to individually optimise two free parameters against the clarity ratings measured in Experiment 1. These parameters were the precision of prior expectations (as measured by their standard deviation), and a perceptual threshold below which the observer rated speech as unclear (Supplementary Fig. 2). The model explained 97.6% of the variance in the group-averaged clarity ratings (Fig. 2e). 99.4% of the variance could be explained by additionally accounting for non-linearities in the effect of the increasing sensory detail on perceptual clarity beyond 16 vocoder channels, but analysis of the Akaike information criterion suggested that this increase in variance explained did not outweigh the loss of parsimony compared to the simpler model (see Supplementary Discussion). The simpler model was therefore retained, but all of the group differences and associations between model outputs and neurophysiology reported in the results that follow remained significant if the complex model were used.

Patients had significantly more precise prior expectations than controls (Wilcoxon rank sum $U(11,11) = 83, p = 0.005$; Fig. 2f). There was a trend towards patients having lower perceptual thresholds ($U(11,11) = 99, p = 0.075$), meaning that patients required less sensory detail to give a clarity rating of 2 or higher, reflecting an appropriate downwards extension of the subjective clarity scale rather than a higher level introspective deficit resulting in patients not being ideal observers of their sensory experience (see ‘Discussion’). The model results confirm that the consequence of degraded neural mechanisms for sensory predictions is not that the brain is unable to use prior knowledge (written cues) to modulate perception, but rather that patients with nfvPPA apply their prior knowledge with greater precision and inflexibility.

**Induced oscillatory dynamics.** To examine the effects of nfvPPA and task manipulations on induced oscillatory activity, we performed a time–frequency analysis of the planar gradiometer data, averaged across sensors (Fig. 6). First, we inspected the neural instantiation of predictions by analysing induced activity during the period following presentation of the written word but before the onset of the auditory stimulus. Based on recent studies in the auditory domain we expected this updating of predictions to manifest as an increase in beta frequency oscillations preceding the onset of the spoken word$^{19}$. This was confirmed by SPM analysis across time–frequency space in our cohort, with a significant increase in beta power (10–28 Hz) for both groups of participants beginning around 800 ms after the onset of the written word, i.e. a~250 ms before the onset of the spoken word (cluster FWE $p = 0.001$, Fig. 6). At the time (992 ms) and frequency (24 Hz) of the peak effect for both groups overall, the single subject magnitude of the induced response correlated significantly with their precision of prior expectations as simulated by our behavioural Bayesian model (Pearson’s $r(19) = -0.52, p = 0.017$; Spearman’s $\rho = -0.54, p = 0.012$). A confirmatory SPM across the whole time window confirmed the group difference implied by this relationship, with patients displaying a single cluster of greater induced beta (20–34 Hz) power from 868 ms (cluster FWE $p = 0.010$). There were no induced effects that were greater in controls than in patients.

Second, we complemented our evoked analysis by assessing oscillatory power during the reconciliation of predictions, i.e. after the onset of the spoken word, averaged across sensors. The results in this section were not altered by subtraction of the condition-averaged evoked waveform subtracted from every trial, confirming that these are true induced responses rather than high-frequency contamination from the evoked responses described previously. Figure 7a illustrates the oscillatory power induced by hearing noise vocoded speech for each group, normalised to the pre-visual stimulus baseline and averaged across the whole brain. Across all conditions, the general pattern was for increased alpha and beta power for the first ~200 ms, followed by a desynchronisation from ~200 ms onwards.
SPM across time–frequency space demonstrated a significant main effect of congruency in both groups separately (Fig. 7a). In controls, this effect peaked at 436 ms at a frequency of 12 Hz (FWE $p < 0.001$), with greater suppression of this response for spoken words that matched prior written text. In patients, a similar effect was observed, also at 12 Hz, but with a later peak at 824 ms (FWE $p < 0.001$).

Across all 21 individuals SPM demonstrated two time–frequency clusters that showed a group by congruency interaction. In the first, extending from 276 to 444 ms between 4 and 24 Hz (peak 300 ms, 16 Hz), controls displayed a greater effect of congruency than patients (cluster FWE $p < 0.001$). In the second, beginning at 680 ms and extending beyond the end of the analysis from 6 to 34 Hz (peak 888 ms, 20 Hz), the interaction was reversed, with patients showing a greater effect of congruency than controls (cluster FWE $p < 0.001$).

The scalp distribution and source localisations of this effect are illustrated in Supplementary Fig. 4: effects were restricted to the left hemisphere and localised to areas around superior temporal gyrus. Both interaction clusters remained significant (FWE $p \leq 0.002$) in a confirmatory analysis in which power was normalised to the pre-auditory stimulus baseline. To illustrate the time-course of this oscillatory dissociation, the data from all three sensor types were restricted to 12–24 Hz, encompassing the interactions in both directions, and the total effect of congruency on power in this band across the whole brain was plotted in Fig. 7b.

To investigate this delay in the beta response in individual patients, single subject time–frequency decompositions were performed and the time taken to reach 80% of the peak overall power contrast between matching and mismatching prior knowledge was defined for each subject. The effect latencies for controls were all tightly clustered between 275 and 400 ms (Fig. 7c). Every single patient was delayed compared to every single control, with a range of 412–1048 ms (Fig. 7d). In the patient group neural response latency was negatively correlated with grey matter volume in our left frontal region of interest ($r = -0.68$, $p = 0.042$; Fig. 7e) but not in our left temporal region of interest ($r = 0.34$, $p = 0.36$; Fig. 7f). There was a trend towards a negative relationship between latency and the standard deviation of prior expectations, though this did not reach significance ($r = -0.37$, $p = 0.1$).

To summarise, all participants show the same congruency-induced reduction in activity in the STG, but nfvPPA patients are delayed in showing this response compared to controls. Thus, the differential response of patients and controls reflects a top-down effect of frontal neurodegeneration on brain responses in posterior regions that remain structurally intact (compare, Figs. 5b and 1e) and that respond normally to bottom-up manipulations of speech clarity (Fig. 3).

Coherence and connectivity during prediction reconciliation.

To determine whether these effects are due to frontal degeneration or fronto-temporal disconnection, we examined coherence and connectivity between the frontal and temporal lobe sources of interest (Fig. 5c) during the 900 ms immediately following the onset of each spoken word. We employed two MEG connectivity analysis methods that give complementary information concerning fronto-temporal dynamics during the reconciliation of predictions: Imaginary Coherence, which is immune to volume conduction effects and source spread as well as differences in power, and Grainger causality. These analyses allow us to be confident that relationships we describe are true reflections of the underlying brain dynamics. Both groups had significant fronto-temporal coherence up to around 25 Hz (Fig. 8a). Coherence in the beta band (13–23 Hz) was significantly stronger in patients than controls (Fig. 8b). Therefore an overall reduction of fronto-temporal connectivity cannot explain our observed differences between nfvPPA patients and controls.

Imaginary coherence does not provide robust indices of directionality, because inter-regional interactions potentially occur over more than one oscillatory cycle. We therefore also examined Granger causal relationships between our sources of interest. This metric allows us to look at the directionality of...
non-zero-lag fronto-temporal interactions while still being relatively robust to volume conduction. Granger causality tests whether information from the past activity of one region can predict future activity in another better than its own past. Highly significant bi-directional Granger causal relationships were observed between temporal and frontal sources (Fig. 8c). To compare these, while avoiding confounds due to differences in signal to noise ratio between regions and between individuals (which can alter the magnitude of Granger Causal relationships\(^{29}\)), we divided the magnitude of each frequency value by the across-frequency mean for each individual and region to create a profile of relative influence for each region at each frequency. This analysis demonstrated significantly stronger temporal to frontal Granger Causal relationships at low frequencies, while frontal to temporal influences were stronger at higher, beta band, frequencies (Fig. 8d). These findings are in agreement with a recent study of MEG connectivity during written language comprehension, which showed that rhythmic information from temporal and parietal lobes was carried at lower frequencies than that from frontal cortex\(^{33}\).

Overall, our finding of increased imaginary coherence in the patient group in a frequency band where frontal to temporal Granger causal influences predominate demonstrates that frontal neurodegeneration increases rather than reduces top-down connectivity from frontal to temporal regions during speech perception.

**Discussion**

The principal finding of this study is that neurodegeneration of the frontal language network results in the delayed neural resolution of predictions in temporal lobe. Conversely, the temporal lobe neural responses to bottom-up manipulations of sensory detail were not delayed. This proves that the resolution of sensory predictions is causally mediated by frontal regions in humans. Our source-space analysis demonstrates that frontal regions are working harder overall in nfvPPA patients. Our finding that, in the patient group, coherence is increased in a frequency band where frontal to temporal influences predominate, suggests that increased fronto-temporal interaction is required to reconcile excessively precise predictions. This view is supported by recent observations that left inferior frontal imaginary coherence is decreased in nfvPPA during the resting state\(^{34}\), confirming that the increased fronto-temporal coherence we observe here reflects the specific engagement of these mechanisms during language perception and not simply a global upregulation. Together, the results of this study resolve a key controversy in speech perception by demonstrating that frontal regions play a core role in reconciling predictions with sensory input during speech perception\(^{28–30}\).

This observed impairment of predictive processing has significant perceptual consequences. Most strikingly, frontal neurodegeneration does not reduce the degree to which the brain employs contextual prior knowledge to guide lower-level speech perception. Rather, through Bayesian modelling, we show that prior knowledge of expected speech content is applied in an overly precise or inflexible fashion, thereby producing a larger-than-normal behavioural effect of prior knowledge on nfvPPA patients’ ratings of speech clarity. In validation of computational

---

**Fig. 7** Analysis of induced responses after the spoken word. **a** Total induced power after spoken word onset and main effect of cue congruency by group. **b** Overall induced power difference between Match and Mismatch conditions in the alpha/beta overlap range. **c** Single subject time-frequency profiles for each control. The time taken to reach 80% of the peak power contrast between Match and Mismatch trials is indicated for each individual by the number below the corresponding abscissa. **d** Single subject time-frequency profiles for each patient. **e** Significant negative correlation between frontal grey matter volume (adjusted for age and total-intracranial volume at the co-ordinates in Fig. 5c) and the time taken to express a congruency contrast (**d**). The grey shaded area indicates the 95% confidence band for the regression line, marked in black. **f** No significant correlation between similarly adjusted superior temporal grey matter volume and effect latency.
and theoretical models of predictive coding\textsuperscript{19, 55, 56}, we demonstrate that the precision of participants’ predictions correlates with the magnitude of induced beta-frequency oscillations, which have recently been shown to correspond temporally and quantitatively to the updating of predictions\textsuperscript{19}.

Previous neuro-imaging evidence has suggested that frontally-mediated top-down predictions during speech perception are able to explain MEG response magnitudes\textsuperscript{33, 40} and fMRI pattern information in the superior temporal gyrus\textsuperscript{32}. Here, we go beyond these associations and provide causal evidence for a functional contribution of frontal networks in supporting top-down predictions by demonstrating that disruption of these networks has the remote effect of delaying the reconciliation of predictions in temporal lobe regions that are anatomically intact, with striking behavioural consequences.

Patients with nfvPPA lacked normal modulation of frontal neural activity by cue congruency, were delayed in engaging frontal regions, and showed greater frontal activity and fronto-temporal coherence than control participants. This is analogous to the observation that elderly controls globally upregulate cognitive control networks that are selectively engaged by younger listeners only when speech is degraded, and thus appear to lack difficulty-related modulation of activity\textsuperscript{57}.

In contrast to these changes, patients displayed normal power of evoked neural responses for analyses combined over sensors and had normal neural responses to changes in sensory detail. These observations provide reassurance that the lack of atrophy in auditory regions of temporal lobe does not mask a microscopic abnormality in auditory neural function. Similarly, it is reassuring that in both groups clarity ratings were similar for mismatching and neutral cues, and the response to questionnaires suggests similar ecological perception of speech in most conditions. This excludes trivial explanations for our behavioural results such as patients becoming confused by mismatches or having an altered approach to subjective rating scales. The group difference in uncued identification of vocoded speech was small, and was accounted for in our Bayesian modelling, by using the results of Experiment 2 to individually define the precision of the sensory input for each subject and number of noise vocoder channels. Thus, our observation of abnormal effects of cue congruency in nfvPPA cannot easily be explained by basic auditory processing deficits, or by higher signal detection thresholds.

Most strikingly, patients displayed a significant delay in the effects of cue congruency in temporal lobe in our sensor space, source space and induced analyses. Under the predictive coding framework, these effects arise because the integration of prediction error is an iterative process, whereby predictions are recursively updated in the light of sensory input to minimise error\textsuperscript{32}. For the patients with nfvPPA, the degraded neural architecture and aberrantly precise predictions might mean that this updating requires more iterations and/or that each iteration takes more time. As a consequence more frontal activity and fronto-temporal coherence is observed, and reconciliation of predictions with sensory signals is delayed.

The behavioural data, Bayesian modelling and neurophysiological results all support the proposal that perceptual predictions

![Image](image-url)
operate within an hierarchical generative network, which for speech perceptions spans auditory, superior temporal, and inferior frontal cortices. Predictive coding is one framework to understand such processing, and makes a number of specific claims about the nature of top-down and bottom-up signals, which we evaluate here.

Firstly, the spatial and temporal pattern of neurophysiological responses in elderly control participants replicated those previously demonstrated in young people, which have been successfully modelled in a predictive coding framework, and which cannot be effectively modelled by sharpening theories of neural representation. Current instantiations of predictive coding models state that each stage of the processing hierarchy passes forwards to the next stage only prediction errors and backwards only predictions. An alternative hypothesis—which might explain some of the present data—is that prior expectations are set up in frontal regions but not fed back to superior temporal regions. When auditory input is received, a process of sensory analysis begins in temporal lobe, with the output fed forwards to frontal regions in real-time, where a matching process occurs. If frontal regions detect that the auditory information matches expectations, they indicate that further processing is unnecessary by feeding back a stop signal to temporal regions. Such a mechanism could also account for our observation of a delayed reduction in superior temporal activity when the text cue is matching. However, this stop-signal hypothesis is unable to account for an increase in top-down, frontal to temporal, connectivity in the patient group. It is also inconsistent with fMRI evidence demonstrating an interaction of prior knowledge and sensory detail in superior temporal representations of degraded speech. These fMRI findings can only be simulated by a computational model in which superior temporal regions represent the discrepancy between predicted and heard speech (i.e., prediction error). In the predictive coding model, the delayed neural effects of cue congruency observed here reflect an iterative process whereby predictions are recursively updated to minimise error; this process operates more slowly in our patient group, and is reflected in greater fronto-temporal coherence.

Most models of prediction and perception, including our own Bayesian modelling, make the assumption that perceptual outcomes represent an ideal observation of peripheral sensation. This might not be the case if individuals hold aberrant beliefs about the fidelity of their sensory input based on differences in previous experience. The results of our Bayesian modelling are inconsistent with the view that our patients are not ideal observers of their sensory experience. If patients with nfvPPA had learnt that their auditory input were unreliable, this could only explain the present data by proposing a dissociation between an underestimation of the precision of their sensory input when reporting perceptual clarity (Experiment 1) and an intact ability to discriminate sensory features when distinguishing alternative vodced words (Experiment 2). This would manifest in our Bayesian modelling as an increase in perceptual threshold, as any given distribution of sensory input would be reported as less clear. In fact, we found that patients did not statistically differ from controls in terms of their perceptual thresholds. Indeed, the trend was towards lower thresholds, which might reflect an appropriate downwards extension of the bottom-end of their perceptual clarity rating scale to reflect the fact that they were slightly less good at identifying vodced speech than controls (Experiment 2), indicating that they had access to slightly less sensory detail during the experiment. Therefore, our behavioural results cannot be accounted for by patients with nfvPAP not being ideal observers of vodced speech.

The hypothesis that beta oscillations represent the instantiation of predictions has existed for some time (see ref. for review). It has been supported by evidence including computational simulations, and empirical observations of backward beta connectivity in speech processing. More recently, direct recordings from human auditory cortex have directly linked beta frequency oscillations to the updating of predictions, based on correlations between observed brain activity and Bayes-optimal predictions generated from presented stimuli. In our cohort, we not only replicate the finding of beta oscillations as a correlate of prediction instantiation, but go further in demonstrating that, irrespective of disease status, the strength of this beta activity across participants relates to the precision of their predictions, as determined by our Bayesian behavioural modelling.

In clinic, patients with nfvPPA often complain of difficulties in hearing speech that are disproportionate to any measured deficits and resistant to hearing aids. We found a dissociation in subjective assessment between understanding speech in noise (which was rated as equally difficult for both groups), and speech in quiet (which only patients rated as difficult). Greater difficulty in quiet environments may seem counterintuitive at first, but it is consistent with the predictive coding hypothesis. Successful perception and comprehension of speech requires continuous updating of predictions based on sentential context and other cues. In a noisy environment, it is beneficial for listeners to rely heavily on these prior predictions as the patients do, because the sensory signal to noise ratio is poor. In a quiet environment, however, this is a suboptimal strategy as greater reliance can and should be placed on more precise or informative sensory inputs. If patients are unable to flexibly adapt the precision of their predictions to quiet listening environments, speech-in-quiet will remain difficult. It might be that globally strong predictions in nfvPPA are an adaptation to their inflexibility, as dysfuctioning networks are forced to choose one prediction strength for all scenarios. If, instead, predictions were globally weakened, this would be beneficial to the perception of speech-in-quiet but detrimental to speech-in-noise, perhaps having a greater overall cost to intelligibility in a dynamic environment (see ref. for similar arguments in motor control).

We probed for subjective clarity ratings around 1050 ms after the onset of the spoken word (Fig. 2a). It could be argued that the abnormal precision of prior expectations in patients might be adaptive to the experimental context. Delayed processing might mean that they have less time for predictions to be enacted before a decision must be made, and therefore stronger predictions are required if they are to have meaningful effects. If so, our finding of increased prior precision in patients might be attenuated if clarity ratings were requested after a longer delay. However, predictions on this slower timescale would be of limited real-world consequence for speech perception and comprehension, because content-containing words such as nouns and verbs are separated by similarly short intervals in typical sentences, the temporal range in which human perception is optimal. Visual cues from lip reading also operate over millisecond timescales and are mediated by similar increases in fronto-temporal functional connectivity to those we demonstrate here.

As well as explaining subjective difficulties with speech perception in nfvPPA, domain-general inflexibility in predictions could account for two other poorly understood behavioural abnormalities in this group. As shown in Supplementary Fig. 5 and Supplementary Results, we replicate the previous observation that deficits in basic auditory processing are overrepresented in patients with nfvPAP. Secondly, agrammatism is a prominent symptom in both nfvPPA and its vascular analogue Broca-type aphasia. Patients are observed to have particular difficulties understanding complex grammatical structures containing
hierarchical structures or the passive voice. These structures are infrequent in daily language, and it would be reasonable to hold a prior expectation for more frequent, subject-oriented

word orders. If patients are less able to flexibly modify this prior on the basis of violations (in this case grammatical cues), it could account for their selective behavioural impairment. This view is consistent with emerging perspectives of linguistic processing as a specialised function of a more general cognitive computational system for complex and flexible thought, based on dynamic functional interactions between inferior frontal and superior temporal cortex. It is empirically supported in our group by recent evidence demonstrating that patients with frontal-limbic aphasias are impaired at detecting violations of ordering relationships in structured auditory sequences, regardless of whether those sequences are constructed of linguistic or non-linguistic stimuli.

Methods

Ethics. All study procedures were approved by the UK National Research Ethics Service. Protocols for MEG and MRI were reviewed by the Suffolk Research Ethics Committee, and for neuropsychological tests outside of the scanner environments by the County Durham & Tees Valley Ethics Committee. All participants had mental capacity and gave informed consent to participation in the study.

Participants. Eleven patients with early nfvPPA were identified according to consensus diagnostic criteria. The diagnosis of degenerative language disorders is complex, and nfvPPA is characterised by an aphasia with prominence of speech and/or agnosia but without problems in single-word comprehension or object knowledge and naming. Particular care was taken to exclude patients with yes/no confusion that would confound behavioural analysis, and to include only those who lacked the lexical difficulties of logopenic and mixed aphasias. This was done in order to select patients most likely to have underlying Tau or TDP-43 related pathology preferentially involving frontal lobes (rather than Alzheimer-type pathology of parietal lobes) and superior temporal gyrus (MNI [−39, −24, −9]) and superior temporal gyrus (MNI [−73, −17, 3]). Frequentist probability of atrophy and Bayesian probability of no atrophy are reported at each of these four locations in results.

To create Fig. 1c, a rendering of the significant regions in each analysis, the DARTEL template images were further warped using the ‘Population to ICBM Registration’ function with the transformation parameters applied to all thresholded statistical maps.

To extract grey matter volume for correlation with the latency of MEG responses (Fig. 7c, f), a full factorial general linear model was constructed with the patients alone, with the patients added, and the time points. The residuals were then used to test for an effect of age and TIV adjusted grey matter volume was extracted at the voxel closest to the MEG regions of interest [left frontal [−62 28]; left temporal [−36 −34 12]]. A secondary SPM analysis with neutral latency entered as a covariate into the model and small volume correction of 8 mm (to match the FWHM Gaussian smoothing kernel) at each location confirmed the result of the post hoc analysis below p < 0.05 were observed at the frontal but not the temporal location.

Modifications to the Sohoglu MEG paradigm. Stimuli and experimental procedures during neuroimaging were closely modelled on a task previously performed to evaluate influences of prior knowledge and sensory degradation in young, healthy listeners (Figs. 1c, 2a). In this task, individuals are presented with a written word, followed 1050 (±50) ms later by a spoken word, which is acoustically degraded using a noise vocoder. After a further delay of 1050 (±50) ms, participants are asked to rate the perceptual clarity of the vocoded word. This allows for a factorial manipulation of two stimulus dimensions that in previous studies have been shown to affect speech perception: (1) the degree of correspondence between the written and spoken words can be modified by presenting text that either matches or mismatches with the speech, (2) the number of distracter items can be manipulated by varying the number of channels in the noise vocoder. 108 trials of each condition were presented across six blocks. Each block contained 18 trials of each combination of vocoder channel number and cue congruency in one of two fixed random orders counterbalanced across groups. To avoid predictability, each stimulus was used at least 216 times, with each word and channel combination used only once as a match pair and once as a mismatch pair) and 108 words only once (in either a match or mismatch pair). The following modifications were made to the Sohoglu et al. paradigm in order to simplify procedures for patients and elderly controls. The number of channels used in the vocoder was doubled to 4/8/16, the range expected to cover the steepest portion of the psychometric response function in older adults. The duration of the visual prime was increased from 200 ms to 500 ms. The resolution of the clarity rating scale was reduced from 1–8 to 1–4, so that a four-button box could be used to indicate responses. Finally, the neutral priming condition was removed to reduce the overall number of trials included in the analysis by a third, while minimising the reduction in the power with which we could test for an effect of prime congruency. 108 trials of each condition were presented across six presentation blocks. A fully crossed 2 × 3 factorial design was employed, with two levels of prime congruency (matching/mismatching) and three levels of sensory detail (4/8/16 vocoder channels). Each spoken word was presented no more than twice to each participant, once with a congruent prime and once with an incongruent prime.

Behavioural data stimuli and procedure. To ensure that we were observing an effect of prediction and that patients were not simply being confused by mismatching written words, experiment 1 was repeated outside the scanner with identical parameters but an additional, neutral, cue condition (Supplementary Fig. 1c), mirroring that of ref. 60. Eighteen trials of each condition were presented in a single block. A second experiment was undertaken to assess participants’ ability to identify vocoded words. In this task, no prior written text was provided. Participants simply heard a noise vocoded word and, 1050 ±50 ms after word onset, were presented with four alternatives, from which they selected the word that they thought was the correct response. The response option “0” (no neighbours in the response array), (2) “Correct” (one correct and one incorrect neighbour “Game”), (3) “Game” (one offset neighbour “Gaze”), or (4) “Then” (no neighbours in the response array). This achieved a factorial experimental design with a 3-level manipulation of
sensory detail fully crossed with a 4-level manipulation of distractor difficulty. There were 90 trials in a single block; at each of the three levels of sensory detail there were ten spoken words in each of cases (1) and (4) above, and five each in (2) and (3). This resulted in 30 sets of four response options each being presented three times and having three of its four members heard during the experiment. Word presentation orders were randomised across participants, but the sensory detail and written word order was fixed. Finally, we exactly replicated a subset of the tasks used by Grube et al.62 to demonstrate peripheral auditory processing deficits in nVPPA. We selected a cross section of tasks on which the patients with nVPPA had displayed particular difficulties, covering a range of processing from simple to complex. These comprised pitch change detection (Grube task P1), 2 Hz and 40 Hz frequency modulation detection (Grube tasks M1 and M2), and dynamic ripple discrimination (Grube task M4)62.

Auditory stimuli were presented in a quiet room through Sennheiser HD250 linear 2 headphones, driven by a Behringer UCA 202 external sound card, and visual stimuli were displayed on a laptop computer screen. Participants responded either by pressing a number on a keyboard (clarity rating outside of MEG) or a button on a custom made response box (all other experiments).

**Behavioural data modelling.** Subjective ratings of clarity were modelled using an hierarchical Bayesian inference approach previously described for data of this type33 (Supplementary Fig. 2). This model exploits the principles of predictive coding, in which perception arises from a combined representation of sensory input and prior beliefs4, 5, 7. It is able to explain both the perceptual benefit of high sensory detail and (3). This resulted in 30 sets of four response options each being presented three times and having three of its four members heard during the experiment. Word presentation orders were randomised across participants, but the sensory detail and written word order was fixed. Finally, we exactly replicated a subset of the tasks used by Grube et al.62 to demonstrate peripheral auditory processing deficits in nVPPA. We selected a cross section of tasks on which the patients with nVPPA had displayed particular difficulties, covering a range of processing from simple to complex. These comprised pitch change detection (Grube task P1), 2 Hz and 40 Hz frequency modulation detection (Grube tasks M1 and M2), and dynamic ripple discrimination (Grube task M4)62.

Auditory stimuli were presented in a quiet room through Sennheiser HD250 linear 2 headphones, driven by a Behringer UCA 202 external sound card, and visual stimuli were displayed on a laptop computer screen. Participants responded either by pressing a number on a keyboard (clarity rating outside of MEG) or a button on a custom made response box (all other experiments).

**MEG and EEG data acquisition and analysis.** An Elekta Neuromag Vectorview System was used to simultaneously acquire magnetic fields from 102 magnetometers and 204 paired planar gradiometers, and electrical potentials from 70 Ag-AgCl scalp electrodes in an EasyCap extended 10–20 system, with additional electrodes providing a nasal reference, a forehead ground, and paired horizontal and vertical electrooculography. All data were digitally sampled at 1 kHz and high-pass filtered above 0.01 Hz. Head shape, EEG electrode locations, and the position of three anatomical fiducial points (nasion, left and right pre-auricular) were measured before scanning with a 3D digitiser (Fastrak Polhemus). The initial impedance of all EEG electrodes was optimised to less than 5 kΩ, and if this could not be achieved in a particular channel, or if it appeared noisy to visual inspection, it was excluded from further analysis.

During data acquisition, the 3D position of five evenly distributed head position indicator (HPI) coils was monitored relative to the MEG sensors (magnetometers and gradiometers). These data were used by Neuromag Maxfilter 2.0, to perform Signal Source Separation73 for motion compensation, and environmental noise suppression.

Subsequent pre-processing and analysis was undertaken in SPM12 (Welcome Trust Centre for Neuroimaging, London, UK), FieldTrip (Donders Institute for Brain, Cognition, and Behavior, Radboud University, Nijmegen, The Netherlands) and EEG lab (Swartz Center for Computational Neuroscience, University of California San Diego), implemented in MATLAB 2013a. Artefact rejection for eye movements and blinks was undertaken by separate independent component analysis to disentangle the contribution of the three sensor types. For MEG data, components were automatically identified that were both significantly temporally correlated with contemporaneous electrooculography data and spatially correlated with separately acquired template data for blinks and eye movements. For EEG data, components were spatially and temporally consistent with eye blinks were automatically identified with ADJUST73. These components were then projected out from the dataset with a translation matrix. Due to a technical difficulty during acquisition, one control subject had no signal recorded from two thirds of their EEG sensors, and one patient had seven sensors that failed quality control—these individuals were excluded from the EEG analysis, but included in MEG and behavioural analyses. For MEG data, the evoked analysis time window was 1–1500 ms relative to speech onset, downsampled to 250 Hz, EEG data referenced, baseline corrected as above, low-pass filtered below 10 Hz, and notch filtered between 50–60 Hz, then epoched. Data were submitted to four separate time frequency decompositions by the Morlet wavelet method: two separate time windows of ~500 to 1500 ms relative to written word and speech onset were examined, with and without pre-subtraction of the condition-averaged waveform from every trial. These were robustly averaged and low-pass filtered below 50 Hz, then resampled to 10 Hz. EEG decomposition parameters were focused for sensitivity to low-mid frequencies, with seven wavelet cycles in a range from 4 to 80 Hz in steps of 2 Hz.
Sensor-space evoked analysis. For each sensor type separately, a flexible factorial design was specified in SPM12, and interrogated across all participants for main effects of prime congruency and clarity. For all sensor types, a scalp position of peak effect was defined where peak FWE $p < 0.01$. The sensor data at this scalp position was then compared across groups at every time point. A significant group×condition interaction was defined as at least seven consecutive timepoints of $p < 0.05$, exceeding the temporal smoothing induced by lowpass filtering at 40 Hz. This approach does not represent double dipping as the location of interest was defined by an orthogonal contrast76–78, and in any case for the effect of congruency (where group×condition interactions were observed with this method), for both the planar gradiometers and the magnetometers the location of peak effect for patients alone was within 2 mm of the conjoint peak effect.

Evoked data source reconstruction. Source inversion methods by the sLORETA algorithm were identical to those employed by Sohoglu et al.80, except that they were undertaken in SPM12 rather than SPM8. It was observed that the time windows of interest defined in healthy young controls were slightly earlier than the main data features in our cohort of more elderly controls, who displayed similar overall profiles to patients with nfPPA (Supplementary Fig. 3). The time windows of interest were therefore slightly lengthened and delayed, to ensure that the main data variance was captured.

The aim of the source data analysis was to localise and explore the brain basis of the group by congruency interaction statistically demonstrated in the sensor space data. While localisation of the main effect of clarity was undertaken across all individuals, and is shown in Fig. 5, in the absence of a group by clarity interaction in sensor space, no further analysis was performed on this condition. From the previous studies in healthy young controls, it was anticipated that significant main effects of congruency would be observed in opposite directions in left frontal regions and left superior temporal gyrus. This was indeed the case, with a small left frontal region being significantly more active across the whole time window with Matching prior information, and a larger region centred on left superior temporal gyrus being significantly more active with Mismatching prior information. These peak locations were defined as voxels of interest, and the source power averaged for each condition at each location within every time window of interest for every individual. Independent, repeated measures ANOVAs were then performed in each time window. Those that demonstrated a statistically significant main effect of group or a group by congruency interaction are illustrated in Fig. 5c (the main effect of congruency was not examined, as this would represent double dipping at these voxels).

Sensor-space induced analysis. The primary analysis of induced data was undertaken in the planar gradiometers because of their superior signal to noise ratio for data of this kind57. Other sensor types were examined secondarily to check for consistency of effect, which was confirmed in all cases. Visual inspection of the time×frequency data at a variety of scalp positions revealed no clear difference in the pattern of effect (although its strength differed, as shown in Supplementary Fig. 4), so all data were averaged across all sensors for statistical comparison. A flexible factorial design was specified in SPM12 for time×frequency data across a time window of ~100 to 1000 ms, and interrogated across all participants for all contrasts of interest (main effects of group, prime congruency and sensory detail, and group×congruency interaction and sensory detail, all pairwise and the three-way interaction). A second, confirmatory, analysis was performed by arranging the time-subtracted evoked data per condition from every trial with identical statistical results, demonstrating that the effects were induced rather than evoked (we make no claims as to whether they are dynamic or structural58).

Induced data source reconstruction. The significant group by condition interactions observed in alpha and beta frequency bands were localised with the ‘Data Analysis in Sensor Space’ toolbox in SPM12. sLORETA was not available in this toolbox, so for closest comparability with the evoked reconstructions, the eLORETA algorithm was used. Reconstructions used time frequency data at the frequency of maximum group×congruency interaction, ±6 Hz. Data were truncated at 50 principal components, to avoid any problems with beamforming after Signal Source Separation, which reduces the number of independent components in the data to around 704, 6. Lead fields were calculated over a window of interest from 350 ms to 900 ms and source reconstructions in three source time windows of equal duration defined by the sensor space group×congruency interaction: 300–450 ms, where controls had a greater main effect of congruency: 450–600 ms, where there was no group×congruency interaction; and 600–750 ms, where patients had a greater main effect of congruency. A flexible factorial design was specified in SPM12, and the group by congruency interaction (already statistically demonstrated across the whole brain) thresholded for visualisation at uncorrected $p < 0.01$.

Coherence and connectivity analyses. The timeseries of the frontal ([$46, 2, 28$] and temporal ([−56, −34, 12]) sources of interest (Fig. 5b) were extracted between 0 and 912 ms after every spoken word using the function spm_eg_inv_extract. The condition-averaged waveform (i.e. the evoked response) in each source was then subtracted from every event to trial result in data with zero-mean and approximate stationarity within the time window of interest. The Fourier spectra were then computed in FieldTrip using multitapers with a ±4 Hz smoothing box. This decomposition was then subjected to separate FieldTrip connectivity analyses with either imaging (i.e. the whole brain) or Granger causality. This same procedure was repeated 1000 times with the trial labels in each region shuffled to create a null distribution. Statistical assessment of the presence of coherence or connectivity at each frequency involved the comparison of the observed data against the null distribution (Fig. 8a, c). Between-group comparisons of imaginary coherence employed unpaired t-tests with unequal variance (the normality assumption was not violated), cluster corrected for multiple comparisons (Fig. 8b). To compare the strength of Granger causal relationships between regions, we first corrected for differences in signal to noise ratio between participants and regions by dividing the magnitude of each frequency value by the across-frequency mean for that individual-region pair. This created a profile of relative influence for each region at each frequency, corrected for overall differences in signal strength. At each frequency, the significance of ‘directionality’ (i.e. temporal to frontal vs frontal to temporal) was assessed with a repeated measures general linear model, and the output corrected for multiple comparisons (Fig. 8d).

Data availability. The processed data that support the findings of this study are available on request from the corresponding author T.E.C. The raw data are not publicly available due to file size, and because participant consent was not obtained for such data sharing, but anonymised data may also be requested for non-commercial academic research purposes. Code for the Bayesian source reconstruction is available from https://github.com/thomascope/Bayesian_Model_Code for the MEG pre-processing and analysis is available from https://github.com/thomascope/VESPA/tree/master/SPM12version/Standalone%20preprocessing%20pipeline.

Received: 23 June 2017 Accepted: 27 October 2017
Published online: 18 December 2017

References
1. von Helmholtz, H. Helmholtz’s Treatise on Physiological Optics, Wisconsin, Vol. 3 (Optical Society of America, 1925).


**Author contributions**
The study was conceptualised by M.H.D. and J.B.R., and designed by T.E.C., E.S., K.P., C.D., M.G., R.P.C. and M.H.D. Data were collected by T.E.C. and J.W., and analysed by T.E.C. with assistance from E.S., W.S. and P.S.J. All authors contributed to interpreting the results and writing the paper.

**Additional information**
Supplementary Information accompanies this paper at 10.1038/s41467-017-01958-7.

Competing interests: The authors declare no competing financial interests

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Artificial grammar learning in vascular and progressive non-fluent aphasias

Thomas E. Copea,b,⁎, Benjamin Wilsonb,1, Holly Robsonc, Rebecca Drinkall, Lauren Deanc, Manon Grube, P. Simon Jonesa, Karalyn Pattersona,d, Timothy D. Griffithsb, James B. Rowea,d, Christopher I. Petkovb

a Department of Clinical Neurosciences, University of Cambridge, UK
b Institute of Neuroscience, Newcastle University, UK
c School of Psychology and Clinical Language Sciences, University of Reading, UK
d Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK

ARTICLE INFO

Keywords:
Aphasia
Grammar
Stroke
Frontotemporal dementia
Implicit learning

ABSTRACT

Patients with non-fluent aphasias display impairments of expressive and receptive grammar. This has been attributed to deficits in processing configurational and hierarchical sequencing relationships. This hypothesis had not been formally tested. It was also controversial whether impairments are specific to language, or reflect domain general deficits in processing structured auditory sequences. Here we used an artificial grammar learning paradigm to compare the abilities of controls to participants with agrammatic aphasia of two different aetiologies: stroke and frontotemporal dementia.

Ten patients with non-fluent variant primary progressive aphasia (nfvPPA), 12 with non-fluent aphasia due to stroke, and 11 controls implicitly learned a novel mixed-complexity artificial grammar designed to assess processing of increasingly complex sequencing relationships. We compared response profiles for otherwise identical sequences of speech tokens (nonsense words) and tone sweeps. In all three groups the ability to detect grammatical violations varied with sequence complexity, with performance improving over time and being better for adjacent than non-adjacent relationships. Patients performed less well than controls overall, and this was related more strongly to aphasia severity than to aetiology. All groups improved with practice and performed well at a control task of detecting oddball nonwords. Crucially, group differences did not interact with sequence complexity, demonstrating that aphasic patients were not disproportionately impaired on complex structures. Hierarchical cluster analysis revealed that response patterns were very similar across all three groups, but very different between the nonsense word and tone tasks, despite identical artificial grammar structures.

Overall, we demonstrate that agrammatic aphasics of two different aetiologies are not disproportionately impaired on complex sequencing relationships, and that the learning of phonological and non-linguistic sequences occurs independently. The similarity of profiles of discriminatory abilities and rule learning across groups suggests that insights from previous studies of implicit sequence learning in vascular aphasia are likely to prove applicable in nfvPPA.

1. Introduction

Aphasia is an impairment of speech and language that often leaves other cognitive and intellectual capacities preserved. Patients with non-fluent aphasias due to frontal lobe damage exhibit significant impairments in grammar (Caramazza and Zurif, 1976; Caplan et al., 1985; Berndt et al., 1996). The grammatical impairments in comprehension and production are highly correlated (Berndt et al., 1983), suggesting that they stem from disruption of core syntactic processes rather than processes such as memory, executive function or motor function (Wilson et al., 2011). The deficits are phenotypically similar in patients with damage due to neurodegeneration (non-fluent variant Primary Progressive Aphasia, nfvPPA) and stroke (‘Broca’s aphasia’), however detailed analysis of speech output has
revealed somewhat differential impairments (Patterson et al., 2006; Thompson et al., 2013). Impairments of receptive abilities have not been compared in similar detail.

Beyond these linguistic deficits, patients with aphasia also display auditory domain general processing deficits that are not specifically related to language (Caramazza and Zurif, 1976; Dominey et al., 2003; Patel et al., 2008; Christiansen et al., 2010; Goll et al., 2010; Grube et al., 2012; Gerannayeh et al., 2014b; Zimmerer et al., 2014a, 2014b; Zimmerer and Varley, 2015; Grube et al., 2016). Such studies have raised the possibility that deficits in structured sound processing may play a prominent role in language disorders, but the nature and extent of these deficits remains unclear. It also remains unclear whether impairments in aphasia are specific to the speech domain (Conway and Pisoni, 2008), or also apply to non-linguistic auditory sequences (Christiansen et al., 2010). One study identified impairments in implicit musical sequence learning in vascular aphasia (Patel et al., 2008), but direct comparisons outside of a musical framework are lacking. If artificial grammar learning tasks tap into domain general (rather than language specific) processes, one might expect rule acquisition to generalise from sequences of nonsense words to identically structured sequences of other sounds, such as tones.

It has been commonly held that grammatical impairments are specific to complex linguistic constructs such as hierarchical relationships and the passive voice (Goodman and Bates, 1997; Grodzinsky, 2000), but there is limited evidence for such dissociations (Zimmerer et al., 2014a, 2014b). By contrast, some studies suggest that the processing of adjacent relationships may be disproportionately impaired by frontal lesions involving motor association cortex (Opitz and Kottz, 2012). Recent studies examining artificial grammar learning in agrammatic aphasia secondary to stroke have focussed on linear sentential structures with varying transitional probabilities (Schuchard and Thompson, 2017). A key outstanding question, therefore, is whether agrammatic aphasia is characterised specifically by deficits for more complex linguistic structures or rather by a more global impairment in processing structured auditory sequences (Berndt, 2000).

Artificial grammar learning tasks are particularly well suited for delineating competence in structured sequence processing, as they focus on ordering relationships in the absence of other cues (e.g., semantics, phonology or pragmatics). They test learning of the rules governing the order in which stimuli occur in a sequence (Reber, 1967). Participants are typically exposed to sequences of stimuli that follow certain rules, so that the ordering relationships between the sequence elements can be learned implicitly. They are then tested with novel sequences that are either consistent with these rules or that violate them in some way, to assess learning. The implicit nature of these tasks allows the testing of a wide range of participants, including patients with aphasia. Unlike natural language tasks, it is possible to present structurally identical sequences comprised of different tokens, for example nonsense words or non-linguistic tone stimuli, to assess the contribution of phonological processing. Finally, artificial grammars with multiple levels of complexity can be used to quantify how well participants are able to learn increasingly complex rules, which may more closely reflect those in natural language grammars (Romberg and Saffran, 2013; Wilson et al., 2015).

The ability to process auditory sequences, even when stimuli are meaningless, is strongly linked with linguistic proficiency (Gómez and Gerken, 2000; Conway and Pisoni, 2008; Conway et al., 2010; Frost et al., 2015). Neuroimaging studies have demonstrated that artificial grammar processing engages a left-lateralised network of frontal, temporal and parietal brain areas similar to the set of regions involved in syntactic operations during natural language tasks (Friederici et al., 2000; Ni et al., 2000; Friederici and Kottz, 2003; Petersson et al., 2004; Forkstam et al., 2006; Friederici et al., 2006; Hickok and Poeppel, 2007; Bahlmann et al., 2008; Makuuchi et al., 2009; Folia et al., 2011; Friederici, 2011; Fedorenko et al., 2012; Petersson et al., 2012a, 2012b) and is associated with developmental language impairment (Evans et al., 2009).

The sequence processing ability of patients with non-fluent aphasia has not been systematically compared across aetiologies. Non-fluent variant Primary Progressive Aphasia (nvPPA), also variously known as Progressive Non-Fluent Aphasia (PNFA), nonfluent/agrammatic Primary Progressive Aphasia (naPPA), and Agrammatic Primary Progressive Aphasia (PPA-G), is an adult onset neurodegenerative aphasia characterised by agrammatism and speech apraxia (Gorno-Tempini et al., 2011). It is in many ways the neurodegenerative equivalent of Broca’s aphasia, though some differences do exist in the pattern of speech output impairment (Patterson et al., 2006). The majority of cases are associated with primary tau pathology but a significant minority have TDP-43 related disease (Kertesz et al., 2005; Josephs et al., 2006; Knibb et al., 2006a, 2006b; Mesulam et al., 2014). nvPPA typically leads to subtle structural neuroimaging changes in left inferior frontal and insular cortex (Gorno-Tempini et al., 2004), which correlate with clinical severity (Rogalski et al., 2011). Chronic non-fluent aphasia due to stroke (Broca’s aphasia) results in a similar clinical phenotype of agrammatism and apraxia of speech. The left frontal tissue damage is stable, with partial clinical improvement over time (Kertesz and McCabe, 1977). The extent and pace of this improvement is variable and depends strongly on the integrity of the underlying white matter (Price et al., 2010; Seghier et al., 2016). Better understanding of the abilities of participants with similar symptoms arising from very different aetiologies could provide valuable insights into the neurobiological underpinnings of domain-general and language-related processes, and inform treatment strategies (Brownsett et al., 2014; Geranmayeh et al., 2014a, 2014b).

In the present study, patients with nvPPA, non-fluent aphasia due to stroke, and matched controls were tested on their implicit learning of a mixed-complexity artificial grammar, combining sequencing relationships of increasing complexity using nonsense words or tones. We aimed to test the following linked hypotheses:

1) Rule acquisition differs when structurally identical sequences are comprised of nonsense words rather than non-linguistic tones.
2) Artificial grammar learning ability is similar in patients with vascular and neurodegenerative aphasia.
3) Grammatical impairments in aphasic patients are disproportionately greater for complex, configurational or hierarchical, sequencing operations.
4) Patients with aphasia can improve their ability to detect grammatical disruptions with repeated implicit training.

2. Methods

2.1. Participants

Three groups of participants were recruited. Demographics of the groups are outlined in Table 1. All patients were right handed. One control was left handed. Thirteen patients with mild to moderate nvPPA were identified from specialist cognitive clinics led by authors JBR and TDG according to consensus diagnostic criteria (Gorno-Tempini et al., 2011). These criteria were strictly applied; particular care was taken to exclude non-fluent patients who had lexical

| Table 1 | Subject demographics. Mean (s.d., range). Age leaving education is reported as it is a better measure of highest scholastic attainment than number of years in study. No individuals were mature students. |
|---|---|---|
| Control | nvPPA | Stroke |
| Number | 11 | 10 | 12 |
| Age | 69 (8, 54–79) | 73 (7, 63–82) | 60 (11, 33–74) |
| Age leaving education | 18 (2, 15–22) | 18 (3, 15–25) | 20 (4, 15–26) |
| Years of musical training | 2 (3, 0–10) | 1 (1, 0–3) | 3 (5, 0–13) |
difficulties, in order to select patients most likely to have underlying Tau or TDP-43 related pathology preferentially involving left frontotemporal lobes, rather than Alzheimer-type pathology of parietal lobes (Rogalski et al., 2011; Rohrer et al., 2012; Sajjadi et al., 2012, 2014; Mandelli et al., 2016). Three patients were excluded on the basis of yes/no response confusion (a common early symptom in nfvPPA that might otherwise have reduced our power to detect language specific effects), resulting in 10 complete nfVP datasets. On the short form of the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass et al., 1983) all patients scored 10/10 for receptive naming, 12/12 for special categories, at least 15/16 for basic word discrimination and at least 9/10 for following complex commands. While nfVP exists on a spectrum, with differing ratios of speech apraxia and agrammatism, all of our patients displayed some degree of impairment of expressive grammar in free speech, and all but two displayed impairment of receptive grammar as measured by the sentence comprehension task on the ‘verb and sentences test’ (VAST) (Bastiaanse et al., 2003) (mean 87.5%, range 70–100%). Similarly, the patients varied in their degree of expressive agrammatism, but none was completely unimpaired. Samples of speech from the participants are available in supplementary materials, and speech profiles are shown in Fig. 1A. BDAE profiles were independently rated by authors TEC, HR and KP. Inter-rater reliability for grammatical form was high, with pairwise Pearson correlations of 0.87, 0.85 and 0.85. Areas of significant grey or white matter loss are shown in Fig. 1B, upper panel.

It is widely recognised that patients with nfVP report difficulties with understanding speech (Goll et al., 2010; Cope et al., 2014; Grube et al., 2016). We asked the patients in this study to complete visual analogue scales assessing their difficulty with ‘Understanding speech in a quiet room’, ‘Telling the direction a sound is coming from’, ‘Understanding speech in a noisy restaurant’, ‘Hearing announcements at a bus or rail station’ as well as ‘How loud do people tell you your TV is?’ Compared to a matched group of controls, patients differed only in reporting more difficulty with ‘Understanding speech in a quiet room’ (p = 0.02). Of the Boston Diagnostic Aphasia Examination (BDAE) subscores, this difficulty was strongly correlated only with ‘Grammatical form’ (r² = 0.778, p < 0.001, Supplementary Figure 1).

Twelve patients with non-fluent aphasia due to left sided stroke were recruited from a volunteer database administered by author HR, supplemented by the identification of incident cases by regional research networks. Recruitment criteria were: a single stroke of at least six months chronicity resulting in at least one month of non-fluent aphasia, with MRI evidence of involvement of either left inferior trigone or operculum. Samples of speech from the participants are available in supplementary materials, and speech profiles (triple marked by authors TEC, HR and KP) are shown in Fig. 1A. On the whole, the stroke group had more severe language impairments than the nfVP group. All had some degree of impairment of grammatical form in free speech, and all but two had impairment of receptive grammar on the VAST (mean 70%, range 40–100%). Lesion overlap maps are shown in Fig. 1B, lower panel.

Care was taken to recruit an appropriate control group. During development of the artificial grammar, extensive piloting developed structures for which learning was least influenced by years of education or performance on global cognitive tests. Nonetheless, it is important to minimise this potential confound by avoiding the use of biased volunteer panels, which tend to preferentially recruit highly educated individuals with supra-normal motivation in research tasks. Therefore, we recruited 8 neurological controls with either chronic inflammatory demyelinating polyneuropathy or multifocal motor neuropathy with conduction block, and three spouses of patients with nfVP, resulting in 11 control datasets. These individuals were chosen to represent a cohort of age-matched individuals with healthy brains and similar levels of habitual neurological contact to the patient groups. All scored normally on the Addenbrookes Cognitive Examination – Revised (ACE-R) (mean 96/100, range 92–99) and Raven’s progressive matrices (mean 47/60, range 37–60).

It was possible to perform pure tone audiometry in all patients with nfVP, 9 of the 12 patients with stroke aphasia and 8 of the 11 neurological controls. This demonstrated that the groups had well matched and age-appropriate auditory acuity (Supplementary Figure 2).

2.2. Stimuli

The Artificial Grammar (AG) used here generates sequences of stimuli from 8 unique elements (Fig. 2A). These sequences are governed by a number of rules of increasing complexity. Rule 1) if a ‘C’ element occurs it must be immediately followed by a ‘D’ element. This represents a simple, invariant linear relationship between two adjacent sequence elements, and will henceforth be referred to as the ‘linear’ rule. Rule 2) all of the ‘A’ elements in the sequence must occur before all of the ‘B’ elements. This is a more complex rule, requiring the participants to recognise a general property of the sequences, and will henceforth be referred to as the ‘configurational’ rule (Zimmerer and Varley, 2015). Rule 3) each ‘A’ element type must be paired with the appropriate ‘B’ elements in embedded relationships (e.g., A₁[A₂A₃] A₄B₅). This complex operation requires tracking both the number and the order of the ‘A’ elements and matching these to the subsequent ‘B’ elements, and is referred to as the ‘hierarchical’ rule.

Sequences consistent with the AG are generated by following any path of arrows from start to end in the illustrated state transition graph (Fig. 2A). Ten consistent sequences were used for the exposure phase of the experiment (Table 2). These sequences were of variable length and contained all of the legal transitions possible with the AG. The remaining subset of consistent sequences generated by the AG were kept for the subsequent testing phase, to allow us to present novel, previously unheard sequences (Table 2). During the testing phase, the participants were presented with 4 repetitions each of 6 consistent and 6 violation sequences, in a pseudo-random order. All test sequences were six elements long, meaning that sequence duration could not be used as a cue by participants. The six violation sequences contained two sequences with violations of each of the three AG rules (Table 2). This design allowed us to identify the specific features of the sequences to which the participants were sensitive.

We tested participants with identically structured sequences of both naturally spoken consonant-vowel-consonant (CVC) nonsense words and non-linguistic tone stimuli. The stimuli were designed to provide acoustic cues to highlight the relationships between some of the key sequencing relationships, as follows. In both the CVC and tone experiments, the ‘A’ and ‘B’ elements fell into distinct acoustic categories. In the CVC experiment the ‘A’ elements all took the form “s-vowel-f” (e.g., “sit”) while the ‘B’ elements were “g-vowel-k” (e.g., “gik”). In the tone experiment the ‘A’ elements were all upwards pitch sweeps while the ‘B’ elements were downward sweeps. Furthermore, the A₁-B₁ relationships that are critical to Rule 3 were highlighted by the presence of the same vowel sounds in the nonsense word experiment (i.e., A₁ and B₁ both contain the central vowel ‘a’) or the tone height in the tone experiment (i.e., both A₁ and B₁ pitch sweeps are centred on the same frequency). To ensure that the participants learned aspects of the AG during the exposure phase, rather than simply responding to acoustical properties of the stimuli, we designed the tone stimuli to avoid linear increases or decreases in pitch in the A₁A₂A₃ or B₃B₂B₁ parts of the sequences. Instead, the centre frequencies of the tone sweeps in such a sequence would be ‘mid-low-high-high-low-mid’. The ‘C’ and ‘D’ elements in the nonsense word experiment were designed to be clearly phonetically distinct from the ‘A’ and ‘B’ stimuli, and in the tone experiments they were continuous pure tones of high or low pitch. Example sequences are available to listen to in the supplementary materials.

The nonsense words were produced by a female speaker, recorded with an Edirol R-09HR (Roland Corp.) sound recorder, and combined into exposure and testing sequences using Matlab (100 ms inter-
stimulus intervals, ISI). The average duration of the nonsense words was 477 ms (standard deviation = 7 ms). The tone stimuli were generated using Matlab. The low tone sweeps were linear sweeps between 100 and 150 Hz (i.e., A2 began at 100 Hz and increased to 150 Hz, B2 began at 150 Hz and decreased to 100 Hz). The middle tone sweeps spanned 200 to 300 Hz and the high tone sweeps spanned 400 to 600 Hz. The C and D stimuli were pure tones at 350 and 800 Hz respectively. The duration of all tones was 450 ms, and these were combined into sequences with ISIs of 100 ms. Stimuli were presented through Sennheiser HD250 linear 2 headphones, driven by either an Edirol UA-4X or Behringer UCA 202 external sound card. The amplitudes of all stimuli were root-mean-square (RMS) balanced, and sequences were initially presented to participants at ~75 dB SPL (calibrated with an XL2 sound level meter, NTI Audio). At the start of the exposure phase, participants were asked if this volume was comfortable and clearly audible and, if not, were allowed to freely adjust the volume to their preference.

Exposure sequences and test sequences for the CVC and tone languages were presented in exactly the same fixed pseudo-random order. There were no differences between the orders of sequences in the CVC and tone runs; only the sound tokens used to represent each element in the artificial grammar differed.

2.3. Procedure

The experimental procedure and instructions given to participants were tightly constrained, to ensure that explicit learning strategies and the effect of receptive language difficulties were minimised. The exact wording of the instruction is included in supplementary materials. Participants were exposed to the CVC language for five minutes.
During this time they were simply instructed to listen to the language and to pay attention to the order of the words (the exact script for the instructions is available as supplementary material). They were then tested by being asked to decide whether 48 individual sequences (Table 2) were correct (i.e. consistent with the artificial grammar) or incorrect (i.e. violated the artificial grammar in some way). Participants were able to express their decision either by pressing a button on a keyboard or custom made response box, or by pointing to yes or no on a piece of paper; whichever they found easiest. At the end of a run, general overall feedback was provided with smiley to sad faces (Wong and Baker, 1988) according to overall percentage correct, along with the performance descriptor ‘Great!’ (> 60%), ‘Well’ (55–60%), ‘OK’ (45–55%), or ‘Badly’ (< 45%). This exposure-test cycle was then repeated in an identical fashion for the tone language. Again, they were explicitly instructed that the important thing was the order of the sounds.

After a short break, participants were then re-exposed to the CVC
language for three minutes, before being re-tested. Sequences for both exposure and testing were presented in a different fixed pseudo-random order on each repetition. At the end of this run, feedback was provided relative to the previous CVC run with faces paired with the descriptors ‘Much Better’ (> 110% of previous score), ‘Better’ (105–110%), ‘Same’ (90–105%), or ‘Worse’ (< 90%). This procedure was then repeated for the tone language.

After a longer break, during which tea and biscuits were provided, participants completed a personal details questionnaire, which included questions about musical training and handedness (all patients were right handed). Patients then undertook the short form of the Boston Diagnostic Aphasia Examination (Kaplan, 1983) and the first half of the sentence comprehension section of the Verbs and Sentences Test to assess receptive grammar (Bastiaanse et al., 2003); controls completed an ACE-R and were tested on matrix reasoning (Raven, 1960), similarly demanding tasks of similar duration. Participants then completed another three minute exposure and test session on the CVC and tone languages, for a total of three testing sessions for each language.

Finally, each participant was re-exposed to the CVC language for three minutes, but the testing session that followed was replaced with an ‘oddball’ task. Participants were told that in this final test the ‘incorrect’ sequences were wrong in a different way, but were not explicitly instructed that they were listening for novel tokens. Where an ordering violation would first have occurred in an ‘incorrect’ sequence, the CVC token was replaced by a novel, previously unheard, oddball element (‘fen’, ‘muz’, ‘rol’, ‘dut’, ‘boz’ or ‘cav’). In this way, we were able to assess whether differential performance on the three grammatical rule types was related to other undesired effects such as stimulus ordering.

All individuals undertook all study procedures on a single day, to ensure that differential patterns of performance consolidation during sleep did not confound our findings. The study procedures took up to four hours, including breaks.

2.4. Stroke lesion mapping

The lesioned area of each brain was manually defined on every slice of each patient's 3T T1-weighted MRI scan in FSL, resulting in a 3D lesion mask. The resulting image was then registered to the standard MNI152 brain using FLIRT (FMRIB's Linear Image Registration Tool (Jenkinson and Smith, 2001); affine transformation model, 12 degrees of freedom). This registration matrix was used to register the patient’s lesion mask to the standard space, from which a standardised lesion volume was computed in Matlab. Regions of interest in the left hemisphere (frontal inferior trigone, frontal inferior operculum, rolandic operculum, putamen and caudate) were identified using the aal atlas in SPM12, and the percentage of each sub-region that was lesioned was extracted for analysis (Supplementary Table 1).

2.5. nfvPPA atrophy mapping

Nine of the patients in the nfvPPA group underwent a 3T volumetric T1 MRI scan. From a database of healthy control scans from the same scanner, an age-matched normative sample of 36 individuals was selected by finding the four nearest-neighbours to each patient in terms of age, excluding duplication (mean age of these controls was 73 years). After segmentation, the nine nfvPPA scans and nine nearest-neighbour controls were used to create a DARTEL template. This was then applied to the remaining 27 controls. Resultant images were normalised to MNI space in SPM12 with an 8 mm smoothing kernel, and separate statistical comparisons were performed for grey and white matter, with total intracranial volume and age as covariates (Fig. 1B).

As for the stroke lesion mapping, regions of interest in the nfvPPA patients were then identified using the aal atlas, normalised to MNI space. As distinct from the lesion analysis in the stroke cases, these regions were defined bilaterally. The grey matter volume in each region was then extracted by applying these regions to each individual's modulated, warped grey matter segmentation and correcting for total intracranial volume.

2.6. Analysis

Performance metrics for analysis were based on signal detection theory (Stanislaw and Todorov, 1999). Standard signal detection measures of discriminability and bias rely on the underlying trial difficulty within a run being constant. In our artificial grammar, it was expected that violations of the linear relationship (complexity level 1) would be easier to detect than configurational violations (complexity level 2), which in turn might be easier to detect than violations of the hierarchical structure (complexity level 3) (see Fig. 2B). To accommodate this, we separately calculated d′ and the non-parametric equivalent A′, based on a comparison of performance on each of the three types of violation sequence to all of the consistent sequences. A single value for bias measures was also calculated based on the combined distribution of violation sequences (Fig. 2B). Hit and false alarm rates of 1 were replaced with (n–0.5)/n (where n is the number of trials), and those of 0 with 0.5/n (Macmillan and Kaplan, 1985; Stanislaw and Todorov, 1999). The non-parametric analogues of these signal detection metrics, A′ for discriminability and β′ for bias, were used for the primary analysis.

All statistical analyses were performed in Matlab R2015b with the Statistics and Machine Learning Toolbox unless otherwise specified. Differences from chance performance in both discriminability measures (A′ and d′) and measures of bias (ln(β), c and β′) were assessed for each group and condition separately using one-sample Wilcoxon signed rank tests (the non-parametric equivalent of the one sample t-test).

The effects of group and rule complexity on discriminability were assessed with three separate repeated measures ANOVA tests (one for each test type: CVC, tones and oddball), with the factor ‘participant number’ nested within ‘group’. This parametric statistical test was employed because there is no appropriate non-parametric test for repeated measures designs of the kind employed here; the Friedman test does not allow multiple groups to be compared. Significant results were explored with post-hoc comparisons of population marginal means.

The degree of learning across exposure-test pairs was assessed by fitting a general linear model in Minitab 17 for the CVC and tone languages. The response variable was discriminability (A′) and the factors were ‘participant number’ (nested within ‘group’), ‘rule type’, and ‘run number’. Significant results were explored with post-hoc Tukey’s range tests.

To assess whether the same rules were learned by participants between task types (CVC vs tones vs oddball), and by extension whether learning of the same artificial grammar was transferrable across token types (CVC vs tones), Spearman correlation matrices were constructed based on performance patterns by sequence for each group. From these, hierarchical cluster analysis was performed to construct dendrograms representing the similarity of performance pattern across test and group, using a ‘farthest neighbour’ linkage method with a data-driven inconsistency coefficient (The Mathworks Inc, 2015).

Exploratory regression analyses were performed in Minitab 17. As there were a large number of variables measured for each subject, stepwise regression was undertaken to determine those variables that best predicted artificial grammar learning. This is an automated process to identify a useful subset of predictors by sequentially adding and removing predictors until an optimal model is obtained. Software default alpha-to-enter and alpha-to-remove values of 0.15 were used, with confirmatory analyses at 0.1/0.15 and 0.1/0.1 yielding identical results. Three separate sets of stepwise regressions were performed; one to explore possible correlations between rule discriminability and neuropsychological and language measures, a second to explore correlations between rule discriminability and stroke lesion site, and a third to explore correlations between rule discriminability and grey
Fig. 3. Group performance on sequence identification. Dashed lines represent chance performance. Error bars represent group-wise standard error of the mean. Controls are in blue, nfvPPA in red and stroke in orange. A) Proportion of correct responses for each testing sequence by group and task. Sequences 1–3 were consistent with the grammar and familiar from the exposure phase, while 4–6 were consistent and novel. Sequences 7–12 contained violations of the types indicated (see Table 2). B) Discriminability of each rule type by group for each language type. C) Overall discriminability by group and run number for the CVC language (improving performance by run represents learning over time). D) Overall discriminability by group and run number for the tone language. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
matter volume in nfvPPA. The potential continuous predictors included in the first model set were Age, Raven’s Progressive Matrix score, years of musical training (which we hypothesised might impact tone language difficulty), sentence comprehension (from the Verbs and Sentences Test), and overall aphasia severity (Aphasia Severity Rating Score). For the second set, the potential continuous predictors were age, the proportion of each region of interest lesioned, total lesion volume, and the number of years since stroke. For the third, potential predictors were age, grey matter volume summed across regions of interest in each hemisphere, and corrected whole brain grey matter volume.

3. Results

Raw performance for each individual sequence is shown in Fig. 3A, and the results of the signal detection theory analysis are illustrated in Fig. 3B. The hierarchical cluster analysis is illustrated in Fig. 4. Performance did not differ between sequences heard during exposure and test phases (Fig. 3A sequences 1–3) and those that were novel during the test phase (Fig. 3A sequences 4–6), so these were collapsed. Results are presented for the non-parametric discrimination measure A′, but the same pattern of findings was present for the parametric equivalent d’ (Supplementary Tables 2 and 3).

For the CVC language, discriminability as measured by A′ showed significant main effects of rule complexity and group, but no group x rule interaction or consistent inter-individual differences (Fig. 3B, Table 3A). Post-hoc comparison of marginal means indicated that the group difference is driven by the control participants performing significantly better than participants in the stroke group (p = 0.0058). Participants with nfvPPA performed at an intermediate level, and were not statistically different from either controls (p = 0.14) or stroke (p = 0.50). Further, all groups performed better at detecting violations of linear grammatical rules than configurational (p = 0.0001) or hierarchical (p = 0.0001), which in turn did not differ (p = 0.99).

The tone language discriminability measured by A′ showed significant main effects of rule complexity and group, but no group x rule interaction or consistent inter-individual differences (Fig. 3B, Table 3A). For this language, post-hoc comparison of marginal means indicated that the group difference is driven by the control participants performing significantly better than participants in the nfvPPA group (p = 0.048). Participants with stroke performed at an intermediate level, and were not statistically different from either controls (p = 0.31) or nfvPPA (p = 0.57). Rather than showing lower performance with increasingly complex rule violations, all groups performed significantly more poorly at detecting violations of the configurational rules than both the linear (p = 0.0009) and hierarchical (p = 0.0001), which in turn did not differ (p = 0.73). Possible reasons for this unexpected pattern are discussed below.

All groups performed well at discriminating the oddball stimuli. For all groups, mean and median discriminability was better than for the CVC or tone languages. Crucially, there was no effect of rule type for the oddball language (Table 3A). This is because violations were no longer based on detection of grammatical rules, but on the detection of novel CVC tokens in various positions within the sequence. This observation reassures us that the effect of rule complexity in the CVC language cannot be explained by the position of the violation within a sequence. There was no significant group difference in performance or group x rule interaction (Table 3A), but there was a strongly significant effect of participant, indicating that individuals within each group differed in their ability to detect the novel non-word tokens. Post-hoc comparison of marginal means was not performed, as there was no rationale from the ANOVA to proceed to this.

From Fig. 3A it can be seen that for some sequences, particularly those which violate the more complex rules, participants in all groups performed at a level below chance (50% correct). This would not be expected if participants were simply guessing for these sequences, but would be expected if the grammatical violation was sufficiently subtle that participants made an active decision that the sequence is not consistent with the grammar. In other words, participants would display bias towards stating that a sequence was consistent with the artificial grammar unless they had evidence otherwise, but it was not known whether this tendency would differ between groups (Haendiges et al., 1996; Dickey et al., 2008). A bias towards yes (consistent) was demonstrated, but there was no group difference in bias or group x metric interaction (Table 3B, Supplementary Figure 3).

Participants, including patients, improved with practice. Results of the general linear model analysis including run number are illustrated in Fig. 3C, D, and Table 3C. For the CVC language, there were main effects of rule type, run number and group, but no interactions between run number and either rule complexity or group. This suggests, (a) that all groups were learning across repeated exposure-test cycles, (b) that the amount of learning over time was not different between the three
rule types, and (c) that the group difference in overall discriminability described above was driven by differences in initial grammatical learning, not by a reduced ability to refine the internal grammatical model through feedback and repeated exposure. Tukey tests confirmed that participants improved on each set of exposure and testing (p < 0.05 in all cases). Mean $A^*$ for the CVC task, but not between tasks within each group. This is confirmed by the cluster analysis. Performance on CVC and tone languages across group was further assessed by a two-sample t-test with unequal variance based on the similarities shown in Fig. 4A. In sample 1 were the six within-language similarities (excluding the diagonal) and in sample 2 were the 9 between-language similarities. This confirmed a highly significant group difference in language similarity; $t(12.5) = 5.2$, p = 0.0002. Identical pair-wise tests between groups across language (blinded to overall ability by non-parametric Spearman rank-order correlation) confirmed that performance profiles did not differ between groups (control vs nfvPPA $t(12.3) = -0.12$, p = 0.90; control vs stroke $t(10.2) = -0.76$, p = 0.46; nfvPPA vs stroke $t(12.2) = -0.01$, p = 0.99).

Stepwise regression analysis between overall measures of discriminability and the behavioural measures listed in the methods yielded no statistically significant predictors for CVC or oddball language performance. If performance on CVC linear rules (where performance was highest) is considered in isolation, overall aphasia severity ($\alpha = 0.002$) was the only significant predictor. The only significant predictor for tone language performance was diagnosis ($\alpha = 0.035$), confirming that patients with nfvPPA performed more poorly than those with stroke, independent of aphasia severity.

Lesion volume and site affected performance in the stroke group. Stepwise regression analysis between overall measures of discriminability for each language and the lesion metrics listed in methods yielded a model for the CVC language including only left putamen ($\alpha = 0.06$); in other words, the ability to detect sequences violations decreased with more severe putaminal lesions. For the oddball language, total lesion volume ($\alpha = 0.016$) and involvement of the left ventral frontal operculum ($\alpha = 0.076$) were included in the model (overall p = 0.023), but in opposite directions: participants with larger lesions were less able to detect oddball CVCs, but this deficit was ameliorated if their lesion had a more anterior distribution. The model yielded a model for the CVC language including only left putamen ($\alpha = 0.002$) was the only significant predictor. The only significant predictor for tone language performance was diagnosis ($\alpha = 0.035$), confirming that patients with nfvPPA performed more poorly than those with stroke, independent of aphasia severity.

Grey matter volume affected performance in the nfvPPA group. The best model for nfvPPA performance on the CVC language included age ($\alpha = 0.002$) and total grey matter volume in the left frontal lobe regions of interest ($\alpha = 0.016$), but not in their right sided equivalents or total grey matter volume. These acted such that performance improved with higher left frontal grey matter volume, and also with age. While it might initially seem counter intuitive that older patients performed better, this is likely to reflect the natural loss of grey matter with age. The model is therefore improved by accounting for the fact that any given value of grey matter volume is relatively more atrophic in a younger individual. There were no statistically significant predictors with grey matter volume in the nfvPPA group for performance on the tone or oddball languages.

4. Discussion

This study successfully used a mixed-complexity artificial grammar learning task with speech sounds and tone stimuli to test aphasic patients with two different aetiologies. The principal observations were that: 1) both healthy individuals and patients with aphasia apply strongly contrasting strategies to assess structured sequences depending on whether the sequences consist of linguistic or non-linguistic auditory tokens; 2) patients with vascular aphasia and nfvPPA show similar patterns of auditory sequence processing impairment compared to controls; 3) aphasic patients are not disproportionately impaired on
more complex auditory sequencing tasks, instead displaying a general impairment in processing structured auditory input; 4) patients with aphasia are capable of implicit learning of this kind through repeated exposure/test cycles. We discuss these results in turn in the following sections.

4.1. Rule acquisition differs when structurally identical sequences are comprised of linguistic or non-linguistic stimuli

In all groups, performance profiles on the CVC language followed the expected pattern of linear relationships being more discriminable than configurational or hierarchical structures. By contrast, participants’ judgments about the tone language did not seem to be based on the abstraction of the intended grammatical rules (Fig. 3A, panel 2). This impression was confirmed by correlation and cluster analyses (Fig. 4). Hierarchical clustering based on non-parametric correlations of single subject performance profiles (Fig. 4B) was clearly able to recover the language learned, but not the group structure. This demonstrates that all groups acquired the same set of rules when making decisions about the CVC language; all that differed between groups was their overall performance. Further, a completely different set of rules were acquired for the tone language, but again this learning profile was almost identical across groups. Therefore, rule acquisition was not transferred between the two languages, and the separation of approach to linguistic and non-linguistic structured sequences was strongly maintained in agrammatic aphasia of either type. This is despite the two languages having an identical structure, and being presented and tested in the same order. This finding cannot be trivially explained by a lower level deficit such as the tone language simply being more difficult, more affected by a reduced fidelity of auditory processing or subject to a higher ‘lapse rate’, which would affect discriminability but could not produce the complete dissociation of response patterns shown here across all groups.

Despite extensive exposure, all groups performed poorly at classifying tone sequence number 1 as consistent with the artificial grammar, and indeed seemed to actively reject it, with control performance for this sequence well below chance. This sequence was comprised entirely of tone sweeps, embedded within a recursive structure. The participants also correctly classified sequence 12, which has similar properties, as inconsistent with the artificial grammar. In this case, good performance on this tone sequence does not reflect an ability to extract the hierarchical rule, but rather a consistent tendency to reject the embedded pattern of tone sweep sequences. Overall, therefore, it does not seem that the tone sequences were assessed for the specific violations of the grammatical rules inherent in the artificial language. Instead, they were judged on the overall ‘feel’ of the sequence in a manner very different to the CVC language but entirely consistent across groups.

We therefore infer that, while both patient groups are impaired in their ability to learn and discriminate sequencing rules for both CVC and tone languages, they maintain the same separation of processing of these languages seen in control participants. Taken together, these results imply that domain specific processes exist for linguistic and non-linguistic structured sequence learning, which are preserved even in the presence of acquired grammatical deficits. These processes might therefore engage different brain networks (Geranmayeh et al., 2014a, 2014b). It is possible that this separation is instantiated by an assessment of phonological ‘well-formedness’ in auditory temporal regions (Obrig et al., 2016).

4.2. Artificial grammar learning in aphasia is similar across aetiologies

A second key observation was that patients with nfvPPA and stroke showed similar patterns of performance for the CVC language. All groups were able to correctly classify the grammatical testing sequences as consistent with the exemplary sequences heard during exposure, and this ability fully generalised from the exposure set to the novel sequences not heard during exposure (Table 2; Fig. 3A). While there was a group difference, with performance in the stroke group being poorer on the CVC task, this effect disappeared when overall aphasia severity was accounted for. This novel result in agrammatic patients with primary progressive aphasia provides the evidence to suggest that existing findings on stroke patients should prove applicable to the progressive aphasias.

The lack of an effect of rule type for the oddball task confirms that the pattern of performance for the CVC language cannot be explained by stimulus-level differences such as the position of the violation within the sequence. The lack of a statistical group difference in the ability to detect oddball CVCs, a task performed at the end of the testing session, suggests that the patients’ impairments in this study are not solely due to generic difficulties with performing psychophysical sequence processing tasks, latent yes/no confusion, difficulties with basic auditory processing or differential effects of fatigue between groups.

The only consistent group difference not accounted for by severity was that patients with nfvPPA performed more poorly than those with stroke on the tone based language. There are a number of possible reasons for this. It might be a consequence of the tone language being more affected by the basic auditory sequence processing deficits previously demonstrated in nfvPPA (Goll et al., 2010; Grube et al., 2016). Alternatively, it might reflect involvement of the right IFC, which was spared in the stroke group (Fig. 1B), and is posited to have a role in prosodic and tonality based judgments. Nonetheless, patients with nfvPPA maintained the same pattern of learning as the other groups (Fig. 4), demonstrating that this deficit is a specific difficulty with processing tonal input rather than a breakdown of the separation of phonological vs tonal structured sequence processing.

4.3. Agrammatic aphasic patients are similarly impaired for both complex and simple sequencing operations

As expected, patients did not perform as well as controls, but the magnitude of this performance deficit did not differ by rule complexity, counter to our initial hypothesis. The results demonstrate that aphasic patients showed a global deficit in sequence processing, rather than a selective impairment on complex sequences. Clinically, patients in both groups make errors in the parsing of more complex syntax, and tend to stick to active, subject-relative structures in their expressive language (Grossman and Moore, 2005). We suggest that this does not reflect a specific deficit in the processing of more complex linguistic structures, but rather that these constructions are simply more difficult and therefore more vulnerable to a global deficit. As well as explaining a clinical symptom, this conclusion is consistent with the functional imaging finding that, in nfvPPA, left inferior frontal cortex activity lacks the normal relationship with syntactic complexity (Wilson et al., 2010); it suggests that the efficiency of IFC is so degraded that even the least complex grammatical structures require maximal neural recruitment. This is analogous to the finding that older adults are no longer able to selectively modulate anterior cingulate cortex in response to increasingly difficult listening environments as they have already fully engaged this region in easy listening conditions (Erb and Obroler, 2013).

It is possible that patients attempt to compensate for this deficit by engaging a wider syntactic processing network involving temporoparietal regions (Schoffield et al., 2012; Blank et al., 2016), but that this is insufficient to compensate for lost language function (Wilson et al., 2016). In stroke, where the IFC is lost entirely, the presence of residual ability could be due to complete reliance on this wider network (Thompson et al., 2010), or the involvement of contralateral IFC. Future functional imaging studies of implicit grammar learning will inform this debate. In any case, our findings of a general impact on sequence processing suggest that the grammatical deficits observed in aphasia, and the recovery from these, might reflect higher level, domain general processes (Caramazza and Zurif, 1976; Dominey et al., 2003; Patel
et al., 2008; Brownsett et al., 2014; Gerammary et al., 2014a, 2014b, in press).

The lack of group by rule complexity interactions is unlikely to be explained by limitations in sample size, as in no case was there even a trend in this direction (all interactions with complexity in Table 3 have p-values > 0.5). Therefore, these results appear to represent a genuine null effect, rather than sub-threshold effects that might become significant with a greater sample size. Moreover, the lack of interaction is also consistent in the tone and CVC experiments, although the response patterns between the two are vastly different. Nor can the results be trivially explained by floor effects in processing the more complex relationships for the following reasons: 1) the nfvPPA group show strikingly parallel behaviour in relation to the control group, consistent with a proportional impairment even on the linear sequencing operation (Fig. 3B, panel 1); 2) the stroke patients may well have reached a floor in performance on the complex sequences in the nonsense word task, but even excluding this group did not cause the group by complexity interaction to approach significance; 3) all groups improved over the three testing runs, and performance improvement was parallel in relation to sequencing complexity (i.e. there was a main effect of run but no run-by-complexity interaction, Figs. 3C); and 4) the tone language showed a very different pattern of results to that for the nonsense words yet, again, there was no evidence of a group-by-complexity interaction.

### 4.4. Patients with aphasia show improved performance over repeated cycles

All three groups demonstrated the same amount of learning across repeated exposure/test cycles of the CVC language; the only difference was in their initial levels of performance (Fig. 3C; Table 3C). This suggests that patients with aphasia were able to update their internal model of the artificial grammar based on feedback and implicit comparison with short periods of exposure. It also suggests that patients did not suffer greater effects of fatigue than controls. By contrast, learning did not occur to the same degree for the tone language (Fig. 3D; Table 3C). It is therefore clear that learning was not transferable between the CVC and tone languages, despite them sharing the same underlying artificial grammar structure.

Together, these findings provide a theoretical basis upon which an exposure-based speech therapy for grammar could be built with the aim of improving subjective difficulty with speech comprehension (Supplementary Figure 1). The results imply that there is potential for improvement from an intensive paradigm based on repeated exposure-test cycles. This could in principle be made home-deliverable and patient-led, an approach that has demonstrable efficacy for improving speech production in similar patient groups (Varley et al., 2016); the tasks employed in this study were automated and computer based. The finding that learning did not generalise across modalities implies that such a therapy would need to use linguistic material, as it would be unlikely to be so well learnt with non-linguistic material or to transfer across domains. In contrast, the finding that patients were able to generalise perfectly from sequences heard during exposure (Fig. 3A, sequences 4–6), to the extent that performance did not differ, suggests that such a therapy might not need to be comprehensive with regards to specific sentence structures. Instead, it is envisaged that a graded programme could be designed, such that training focuses initially on those structures that are having most frequent impact on speech comprehension. Clearly our study does not provide evidence that such a therapy would be more efficacious than existing methods for addressing syntactic deficits after stroke, nor do we have any evidence of how well our strategies would work within an already-learned but now-impaired natural language. Indeed, a recent small study of nine patients who had chronic agrammatic aphasia secondary to stroke suggests that implicit learning alone (on a visuo-motor serial reaction time task) does not necessarily translate into improved real-world performance (Schuchard et al., 2017). Larger scale trials and assessment with other tasks are clearly required in more acute disease cohorts. As our method relies on exposure-based implicit learning rather than explicit instruction, future therapeutic trials could follow recent trends for patient-led practice in the home environment, increasingly with support from internet-based resources (Rogalski et al., 2016). This has the potential for providing the benefits of an intensive approach (Bhogal et al., 2003; Brady et al., 2016; Breitenstein et al., 2017) without the resource constraints that limit the frequency and therefore efficacy of traditional irregular, face-to-face instruction (Sarno et al., 1970; Lincoln et al., 1984).

### 4.5. Brain behaviour relationships

Our use of stepwise regression, to assess associations between brain structure and function, should be seen as exploratory, since the study was powered only to detect strong effect sizes. Nonetheless, significant associations between brain structure and behavioural performance within the patient groups were observed. The nfvPPA group demonstrated significant grey and white matter loss in frontal regions bilaterally. In keeping with previous studies of expressive grammar (Rogalski et al., 2011), the discriminability of grammatical structure in the CVC language correlated with age-corrected loss of volume only in left frontal regions (but not similar right-sided regions or total corrected grey matter volume). No such association was found for the tone language or the oddball task.

In the stroke group there was catastrophic loss of fronto-temporal regions in the dominant hemisphere, but complete contralateral preservation. For the CVC language, the model included only the putamen. The putamen is known from the functional imaging literature to be important in implicit sequence detection and learning in healthy individuals (Grafton et al., 1995; Rauch et al., 1995; Rauch et al., 1997; Grahn and Rowe, 2009). No such relationships were significant for the tone task. Performance on the oddball sound detection task was predicted by overall lesion volume and a weaker, opposite, effect of the extent of involvement of the most frontal region analysed. This suggests that more anterior lesion locations were less deleterious to CVC oddball detection than those located closer to auditory cortex in temporal lobe.

Similar performance in a relatively bilateral disease (nfvPPA) and one so clearly unilateral (stroke) immediately poses the question of which components of a broader language network are recruited to underpin the demonstrated learning over repeated exposure-test cycles (Crinion and Price, 2005). Thus, there is clear scope for a functional imaging study to be conducted in these patient groups, the results of which could complement the development of grammar-based speech therapy.

### 5. Conclusion

In this paper we reconcile a controversy in the literature regarding the effects of structural complexity on receptive grammar in the frontal aphasias. We demonstrate that, while the patients found complex, non-adjacent structuring relationships more difficult to acquire, this did not represent a disproportionate impairment; aphasia resulted in a similar performance penalty for adjacent relationships. We also provide insights into the language-specificity of artificial grammar learning by demonstrating that humans learn otherwise identical linguistic and non-linguistic structured sequences entirely separately, even if the neural architecture underlying this learning is disrupted. Our direct comparison of two patient groups suggests that previous findings regarding implicit sequence learning in stroke aphasia are likely to prove transferrable to nfvPPA. Finally, the ability of both patient groups to learn an artificial grammar as demonstrated here provides a rationale and approach for future trials of implicit, exposure-based approaches to rehabilitation of agrammatism in non-fluent aphasia.
Acknowledgements

The Association of British Neurologists, Patrick Berthoud Charitable trust and National Institute for Health Research support TEC. The Wellcome Trust supports BW (WT101918), TDG (WT109644MA), JBR (WT103388) and CP (WT10961MA). The Stroke Association supports HR (TSA 2012/02). Further study support was received from the National Institute of Health Research’s Biomedical Research Centre (Cambridge) and Biomedical Research Units in Dementia (Cambridge, Newcastle).

Appendix A. Supplementary Information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia.2017.08.022

References

Crinion, J., Price, C.J., 2005. Right anterior superior temporal activation predicts audi-


