Neural compensatory mechanisms that support the retention and recovery of language function post stroke

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2021

James D. Stefaniak

School of Biological Sciences

Division of Neuroscience and Experimental Psychology

List of contents

The page numbers for Chapter Three correspond to the pagination of the print version of the published article.

Word count = 40998 (excluding material before Chapter 1 and Appendices)

List of tables

List of figures

The page numbers of the Figures in Chapter Three correspond to the pagination of the print version of the published article.

List of boxes

Abstract

Aphasia is a prevalent and debilitating consequence of stroke that often persists chronically. The overarching purpose of this thesis was to better understand the neural compensatory mechanisms that take place to minimise language dysfunction, and promote language recovery, in post-stroke aphasia. This thesis proposes that existing theories of aphasia recovery can be conceptualised as specific examples of two more fundamental principles, degeneracy and variable neurodisplacement (Chapter 2). 'Degeneracy' predicts that upregulated compensatory regions should not be engaged during language in health, while 'variable neurodisplacement' predicts that such regions might be downregulated during health to save resources but upregulated when task difficulty is increased in controls. The thesis then investigates the neural regions that are functionally involved in language recovery post-stroke. Chapter 3 reports an Activation Likelihood Estimation meta-analysis of coordinate-based language functional neuroimaging studies. The language network is bilateral in aphasia and controls. Regions of the right anterior insula and inferior frontal gyrus were more likely to be activated during language in aphasia than controls, overlap with the Multiple Demand network and were more likely to be activated during higher than lower demand language tasks, consistent with enhanced utilisation of spare capacity within right hemisphere executive regions via variable neurodisplacement. Unexpectedly, Chapter 3 found that multiple undamaged midline and right hemisphere regions were less likely to be activated during language in aphasia than controls, consistent with functional diaschisis. Chapter 4 reports one of the first language multivariate pattern analysis functional imaging experiments in post-stroke aphasia and controls. Chapter 4 suggests the existence of a novel form of 'information diaschisis' in which having a stroke is associated with lower language information processing in a bilateral set of undamaged, predominantly domain-general regions which in turn is associated with, and might contribute to, language impairment post-stroke. Chapter 5 analysed longitudinal functional imaging and neuropsychological data from participants with aphasia post-stroke. Language profiles were multidimensional at 2 weeks post-stroke and could be represented by three orthogonal components representing fluency, semantic/executive function and phonology. Different language components had uncorrelated recovery trajectories that were associated with changing activation in different neural regions during aphasia recovery. This provides insights into the multidimensional nature of aphasia recovery and suggests that future clinical rehabilitation trials should take a 'personalised medicine' approach that accounts for each patient's specific language profile when deciding targets for non-invasive brain stimulation.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

Dedication

I dedicate this doctoral thesis to Reem Alyahya, Pauline Stefaniak and Alex Stefaniak

Acknowledgements

I thank and praise God for blessing me with the ability to complete this work. My heartfelt gratitude goes to Reem, for believing in me, supporting me and changing my life in every way. My sincerest thanks go to my parents. My thanks go to my PhD supervisors, Professor Lambon Ralph and Professor Griffiths.

The author

I was born and raised in Coventry before starting medicine at the University of Cambridge in 2008. In 2011 I graduated with a BA degree (Part II Physiology, Development and Neuroscience, triple first). I transferred to the University of Oxford for my clinical years and graduated with a BM BCh (with distinction) in 2014. I worked on an Academic Foundation Programme in Clinical Neurosciences at the University of Cambridge between 2014 and 2016. I then took up an Academic Clinical Fellowship in Neurology at the University of Manchester in 2016. I obtained Membership of the Royal College of Physicians in 2017. In 2017 I started a Wellcome Clinical Research Training Fellowship and began a PhD in Cognitive Neurosciences at the University of Manchester. I started as a visiting student at the MRC Cognition and Brain Sciences Unit of the University of Cambridge in 2018. In 2020 I began an Academic Clinical Fellowship in Psychiatry at the University of Cambridge. I would not have been able to complete this without the support of Reem Alyahya.

Chapter 1

General Introduction and Thesis Overview

General introduction and thesis overview

This introductory chapter has three sections. First, the overarching research themes of this thesis are outlined. Second, an overview of the structure of this thesis is provided. Third, contributions of other authors to this work are acknowledged. A detailed background on post-stroke aphasia is covered in Chapter 2 and not provided in this introductory chapter.

Research Themes

Aphasia is a prevalent (Engelter et al., 2006) and debilitating consequence of stroke that is associated with increased care costs (Ellis et al., 2012), functional dependence (Boehme et al., 2016) and death (Tsouli et al., 2009). Although acutely resolving biological factors such as reperfusion of key language regions (Hillis & Heidler, 2002) might contribute to language recovery in the acute phase post-stroke, significant recovery of function can occur between the subacute (several weeks post-stroke) and chronic phases that is not associated with corresponding ‘macroscopic’ changes in brain structure or perfusion, and which is thought to be mediated by changes in function of surviving neural regions through what we call 'neural compensatory mechanisms' (Stefaniak et al., 2020). Nevertheless, language recovery can be incomplete and aphasia often persists into the chronic phase (Maas et al., 2012; Pedersen et al., 1995; Wade et al., 1986). The overarching purpose of this thesis was therefore to better understand the neural compensatory mechanisms that take place to minimise language dysfunction, and promote language recovery, in post-stroke aphasia. Only by understanding such compensatory mechanisms, and their neural basis, will we be able to develop rational, neurobiologically-informed therapeutic targets for non-invasive brain stimulation (Bucur & Papagno, 2019; Ren et al., 2014) and other rehabilitation strategies (Conroy et al., 2018; Woodhead et al., 2017), or enable the rational development of neuroimaging-based biomarkers to predict clinical outcome.

This thesis has three overarching themes.

The first theme argues that there are a confusing multitude of existing theories of aphasia recovery that are not placed within an overarching framework; these theories have tended to be verbal descriptions of the neural regions involved rather than theories of the underlying mechanisms themselves, and as such are frequently incapable of being implemented in computational models or making specific predictions that are refutable through empirical observation. Furthermore, such theories have occasionally relied on counterintuitive post-hoc explanations to account for empirical data. For instance, 'right hemisphere' theories have

existed since the nineteenth century (Finger et al., 2003) and posit that the right hemisphere takes on a new or enhanced role during language post-stroke. This, by itself, is not a mechanism but a verbal description of the role for this neuroanatomical region. The current literature contains opposing views about the utility of the right hemisphere in aphasia recovery, with some studies finding its involvement during language to be positive (Crinion & Price, 2005; Robson et al., 2014) but others finding it negative (Postman-Caucheteux et al., 2010; Szaflarski et al., 2013; Tyler et al., 2010). Both positive and negative findings have been accounted for by claiming, for positive associations, that the right hemisphere contains homologous language regions that are quiescent in health but engaged following stroke (Finger et al., 2003; Grafman, 2000); and, for negative associations, that such upregulated homologous regions inhibit the 'primary' left hemisphere language regions through 'transcallosal inhibition' (Heiss & Thiel, 2006; Thiel et al., 2013). This is despite there being no mechanistic explanation for how 'unused' language representations in quiescent regions might resist being repurposed for active use in other cognitive functions (Stefaniak et al., 2020), and no empirical evidence that transcallosal inhibition occurs between language regions (Stefaniak et al., 2020). Other theories include the importance of 'perilesional' regions (Rosen et al., 2000), or the role of domain-general executive regions in the right hemisphere (Geranmayeh et al., 2014). This first theme therefore argues that there is a pressing need to define a coherent mechanistic framework within which the multitude of existing recovery hypotheses can be conceptualised and from which specific predictions can be made that are refutable through empirical observation.

The second theme argues that we need to determine the neural regions that are functionally involved in language and its recovery post-stroke in order to provide empirical evidence for or against the proposed mechanisms of aphasia recovery. This is based on the assumption, outlined above, that increased utilisation of surviving neural regions contributes to aphasia recovery (Murphy & Corbett, 2009; Stefaniak et al., 2020; Turkeltaub et al., 2011), and that increased utilisation should be observable through *in vivo* language functional neuroimaging studies comparing patients with aphasia to matched controls. Functional involvement during language has been described almost exclusively in terms of mass univariate activation during an in-scanner language task (Turkeltaub et al., 2011). However, activation by itself is probably insufficient to contribute to the performance of a cognitive task; rather, activation must help process task-relevant information to contribute to behaviour (Chang & Lambon Ralph, 2020). More modern and advanced 'information-based' functional brain imaging approaches (Kriegeskorte et al., 2006), collectively termed multivariate pattern analysis

(Haynes, 2015), assess whether the distributed activity pattern across multiple voxels in a region can be used to 'decode' which of two cognitive states a participant is in. Such information-based imaging approaches might therefore be better at identifying regions that functionally contribute to language performance in post-stroke aphasia than traditional activation-based approaches. However, very little research has been done using multivariate pattern analysis in aphasia recovery to date (Fischer-Baum et al., 2017; Lee et al., 2017; Li et al., 2021). Relatedly, longitudinal functional imaging studies can provide powerful evidence to adjudicate between which neural compensatory mechanisms occur *in vivo* by identifying how the regions functionally involved in language change over time, as well as by controlling for inter-patient variability in pre-morbid and post-stroke language performance. However, only a handful of studies have performed longitudinal functional imaging during aphasia recovery to date (Cardebat et al., 2003; Geranmayeh et al., 2017; Heiss et al., 1999; Mattioli et al., 2014; Nenert et al., 2018; Stockert et al., 2020; van Oers et al., 2018), particularly during the first few months post-stroke when language recovery occurs most rapidly (Pedersen et al., 1995). This second theme therefore argues that longitudinal functional neuroimaging data, as well as multivariate information-based analyses, are needed to determine the neural regions that are functionally involved in language and its recovery post-stroke.

The third theme argues that language is multidimensional and that distinct language dimensions might rely on different recovery mechanisms or changing functional involvement in distinct underlying neural regions post-stroke. It is increasingly recognised that language is not a single, homogenous cognitive function but instead reflects graded variations along multiple underlying dimensions that have at least partially distinct neural substrates in both the acute (Kummerer et al., 2013) and chronic phase post-stroke (R. Alyahya et al., 2020; Mementi et al., 2011; Mirman et al., 2015). However, previous functional neuroimaging studies in post-stroke aphasia have tended to associate functional involvement during language with either a single behavioural score (Cardebat et al., 2003; Stockert et al., 2020) or with a few language scores that cover a limited part of the full, multidimensional language profile (Mattioli et al., 2014; Nenert et al., 2018; van Oers et al., 2018). The third theme argues that we need to determine the neural regions that are functionally involved in each of the multiple dimensions underlying language in order to obtain a more complete understanding of the neural compensatory mechanisms subserving aphasia recovery.

Outline of Thesis Chapters

This PhD thesis is presented in the journal format, in which the core Chapters (2 to 5) are written in the style of journal papers – which have either already been published (Chapters 2-3) or have been submitted for publication (Chapters 4-5). As a result, there is inevitably some repetition between each of these self-contained Chapters in order to ensure that each publication is self-complete.

This **Introductory Chapter 1** has set out the research themes of this thesis and outlined the structure of the remaining chapters.

Chapter 2 contains a literature review, which has been published: Stefaniak, J. D., Halai, A. D., & Lambon Ralph, M. A. (2020). The neural and neurocomputational bases of recovery from post-stroke aphasia. *Nat Rev Neurol*, 16, 43 - 55. The print version of the published article has been used. Chapter 2 addresses the first research theme by attempting to define a coherent mechanistic framework within which the multitude of existing recovery hypotheses can be conceptualised.

Chapter 3 includes the first empirical study, which has been published open access: Stefaniak, J. D., Alyahya, R. S. W., & Lambon Ralph, M. A. (2021). Language networks in aphasia and health: A 1000 participant activation likelihood estimation meta-analysis. *NeuroImage*, 233, 117960. The print version of the published article has been used. Chapter 3 addresses the second research theme by presenting a systematic review and Activation Likelihood Estimation meta-analysis of coordinate-based language functional neuroimaging data in post-stroke aphasia and controls. Chapter 3 advances our understanding of the neural basis of aphasia recovery using mass univariate activation as a proxy for functional involvement during language.

Chapter 4 includes the second empirical study, 'Information diaschisis within language and domain-general regions in chronic post-stroke aphasia'. Chapter 4 addresses the second research theme by presenting one of the first language multivariate pattern analysis functional magnetic resonance imaging studies in participants with chronic post-stroke aphasia and controls. Chapter 4 advances our understanding of the neural basis of aphasia recovery using multivariate information processing as a proxy for functional involvement during language.

Chapter 5 includes the third empirical study, 'The multidimensional nature of aphasia

recovery post stroke'. Chapter 5 addresses the second and third research themes by analysing longitudinal data from a cohort of individuals with post-stroke aphasia who underwent functional neuroimaging and behavioural testing at ~2 weeks and ~4 months post-stroke. This study examined whether distinct underlying language dimensions in the subacute stage post-stroke recover together in a homogenous manner and have recovery trajectories that relate to changing activation in distinct or overlapping underlying brain regions. Chapter 5 advances out understanding of the neural basis of recovery of the multiple dimensions underlying language and thus provides a more complete view of the neural compensatory mechanisms subserving aphasia recovery post-stroke.

Chapter 6 provides a general discussion of the thesis by drawing together its results in the light of the research themes discussed in this introductory chapter; outlining its clinical implications; discussing its limitations; and suggesting directions for future research.

Acknowledgement of authors' contributions

Chapter 5 reanalysed data from a separate study that had been independently designed and collected by Dr Fatemeh Geranmayeh and had previously contributed towards separate publications (Geranmayeh et al., 2017; Geranmayeh et al., 2016). Dr Fatemeh Geranmayeh independently designed and collected the neuroimaging and behavioural data used for Chapter 5 and was involved in the conceptualisation and writing of Chapter 5. I was involved in the conceptualisation, design, analysis and write-up for the data reanalysis of Chapter 5. I was significantly involved in the conceptualisation, study design, data collection, analysis and write-up of all other Chapters in this thesis. I recruited all stroke patients and healthy controls and performed a majority of behavioural data collection and imaging data collection for Chapter 4. Professor Matthew Lambon Ralph and Professor Tim Griffiths supervised the work in this thesis. Professor Lambon Ralph was involved in the conceptualisation of Chapters 2-5 and writing of Chapters 2, 3 and 5. Dr Reem Alyahya provided invaluable advice regarding the design, data collection, analysis and write-up of Chapters 1-6, and in particular the conceptualisation, design, analysis, write-up and data visualisation of Chapters 3 and 4. Ajay Halai was involved in the conceptualisation and writing of Chapter 2 and did some of the image data preprocessing and a minor amount of data collection for Chapter 4. Francesca Branzi, Gina Humphries and Grace Rice contributed to discussions regarding the study design for Chapter 4. Alan Archer-Boyd helped with auditory stimulus recording and

preparation and MRI scanner setup for Chapter 4. The support staff at the CBU - in particular, Gary Chandler - were very helpful with the technical set up for data collection of Chapter 4. We would like to thank Professor Cathy Price and the PLORAS team for referring patients to our study in Chapter 4. We would like to thank Daniel Mitchell for advice regarding multivariate pattern analysis in Chapter 4. We would like to thank Sonia Brownsett, Fatemeh Geranmayeh, Glyn Hallam, Jessica Hodgson Narges Radman, Michael Mouthon, Alex Leff, Stephen Wilson, Yuan Tao, Erin Meier, and Elena Barbieri for kindly sharing unpublished coordinates relating to their published studies for use in the meta-analysis in Chapter 3.

Chapter 2

The neural and neurocomputational bases of recovery from post-stroke aphasia

Published paper: Stefaniak, J. D., Halai, A. D., & Lambon Ralph, M. A. (2020). The neural and neurocomputational bases of recovery from post-stroke aphasia. *Nat Rev Neurol*, 16, 43 - 55.

The neural and neurocomputational bases of recovery from post-stroke aphasia

*James D. Stefaniak[1,2], Ajay D. Halai[2] and Matthew A. Lambon Ralph[2]**

Abstract | Language impairment, or aphasia, is a disabling symptom that affects at least one third of individuals after stroke. Some affected individuals will spontaneously recover partial language function. However, despite a growing number of investigations, our understanding of how and why this recovery occurs is very limited. This Review proposes that existing hypotheses about language recovery after stroke can be conceptualized as specific examples of two fundamental principles. The first principle, degeneracy, dictates that different neural networks are able to adapt to perform similar cognitive functions, which would enable the brain to compensate for damage to any individual network. The second principle, variable neuro-displacement, dictates that there is spare capacity within or between neural networks, which, to save energy, is not used under standard levels of performance demand, but can be engaged under certain situations. These two principles are not mutually exclusive and might involve neural networks in both hemispheres. Most existing hypotheses are descriptive and lack a clear mechanistic account or concrete experimental evidence. Therefore, a better neurocomputational, mechanistic understanding of language recovery is required to inform research into new therapeutic interventions.

Neurocomputational
In a neurocomputational model, the structure or function is constrained by neurobiological or neuroanatomical characteristics.

Conduction aphasia
A type of acquired language deficit in which individuals have relatively preserved comprehension but impaired repetition and phonologically disrupted fluent speech.

[1]*Division of Neuroscience and Experimental Psychology, School of Biological Sciences, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK.*

[2]*MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK.*

**e-mail: matt.lambon-ralph@mrc-cbu.cam.ac.uk*

https://doi.org/10.1038/s41582-019-0282-1

'Aphasia' refers to the collection of acquired receptive and expressive language deficits that arise in many acute and progressive neurological diseases or following neurosurgery but are most frequently observed following left middle cerebral artery stroke. Post-stroke aphasia is heterogeneous and was classically defined by categorical subtypes; for example, Broca aphasia[1], comprising non-fluent, agrammatic expressive deficits, or Wernicke aphasia[2], defined as compromised comprehension with expressive and receptive phonological impairments. In recent years, graded multidimensional variations of aphasia have been described that reflect the status of a limited number of interacting neurocomputational components — for example, phonology, semantics, speech fluency and executive skills — in the affected individual[3–6]. Globally, more than ten million new cases of stroke are reported each year[7] and at least one third of the affected individuals will have symptoms of aphasia[8]. Aphasia adds substantial costs to the acute[9] and chronic[10] care of individuals with stroke, and is an independent predictor of subsequent functional dependence and death[11].

Many individuals with post-stroke aphasia exhibit some spontaneous recovery of language function[12]. The degree of recovery differs between patients, and the rate of improvement slows with time[13,14]. The classic view is that a plateau of language function is reached within 6–9 months of stroke[15]; however, small changes (both positive and negative)[16] can occur at later stages[17]. During recovery, both the subtype and the severity of aphasia can change over time; for example, individuals might progress from Wernicke aphasia to conduction aphasia, to anomic aphasia, to 'recovered'[18], although 'recovered' individuals might have mild residual impairments that would be detected by more sensitive assessments[19]. Nevertheless, in at least half of patients some form of aphasia persists into the chronic phase[20]. Therefore, we need to uncover the cognitive and neural mechanistic bases of recovery from aphasia in order to establish pharmacological, speech and language therapy (SLT) or non-invasive brain stimulation (NIBS) interventions that maximize spontaneous recovery.

To date, theories of the mechanisms underlying recovery from aphasia have tended to be anatomical descriptions of the regions involved, rather than computationally or mathematically implemented accounts. In this Review, we argue that the existing theories of recovery from aphasia can be coalesced into two principles: degeneracy and variable neuro-displacement. We outline some of the contemporary methods used to measure language performance and its neural bases, as well as conceptual and computational models of healthy language processing that are needed to understand how

Key points

- The mechanisms underlying recovery from post-stroke aphasia can be conceptualized as the engagement of degenerate networks or the use of spare capacity within or between networks via variable neuro-displacement.
- Degenerate networks are not involved in the language task in the premorbid state, but can be engaged for that task after damage, either immediately or following experience-dependent plasticity.
- Degenerate networks might include quiescent regions in the right hemisphere, the undamaged ventral or dorsal language pathway, or regions that supported a non-language activity before stroke.
- The use of spare capacity within or between neural networks could be downregulated to save energy under standard levels of performance demand but upregulated when performance demand increases, for example when healthy individuals are performing a difficult task or in individuals after brain damage.
- Spare capacity that might contribute to recovery from post-stroke aphasia includes the unaffected regions of damaged neural networks, or undamaged networks that perform other language-specific or domain-general executive functions.
- Most theories of recovery from post-stroke aphasia are descriptive and lack concrete experimental evidence; a better understanding of the mechanisms underlying recovery, preferably in the form of computationally implemented models, is needed and the resultant mechanistic accounts will aid the design of therapeutic interventions.

language networks might change to support recovery from aphasia. This background is followed by a detailed discussion of existing theories of the mechanisms underlying recovery from aphasia. For each theory, we highlight the supporting empirical evidence and describe any relevant computationally implemented mechanistic models. We conclude with the key implications for future research into therapeutic interventions.

Types of data on recovery from aphasia

Two main types of evidence, behavioural and neuroimaging, have advanced our understanding of the cognitive and neural mechanisms of recovery from aphasia (BOX 1). Behavioural assessments of language and, in recent years, non-language cognitive abilities[21] have substantially advanced our understanding of the cognitive bases of recovery from aphasia. Most large-scale studies have used cross-sectional data[5,6,22], which do not account for the premorbid or acute status of the participants. However, a handful of logistically challenging longitudinal investigations have now been published[23–25]. The types of behavioural assessments that are used to study aphasia have also evolved since the classical nineteenth century studies of Broca, Wernicke and their contemporaries, moving from classifications of aphasia subtypes[26] to targeted assessments of specific language activities[3]. Longitudinal changes in language performance provide key insights into the cognitive basis of recovery; for example, an improvement in comprehension but not speech production would suggest recovery of the semantic but not phonological or motor systems. Combining these behavioural profiles with neuroimaging gives clues to the neural basis of the initial impairment and the capacity for recovery. Neuroimaging (for example, MRI, magnetoencephalography, EEG and PET) can now be used to extract structural, functional and connectivity data. Each type of neuroimaging has intrinsic advantages and disadvantages, which might be overcome by adopting a multimodal imaging approach (BOX 1).

Healthy language function

Before we can consider how the language system might change to support recovery from aphasia, we need to understand healthy language functions and their neural bases. A complete overview of our knowledge of healthy language function is beyond the scope of this Review, but some key points are considered in this section.

Language is not a singular neurocognitive function: language activities are diverse, including comprehension, speaking, reading, writing and repeating, and are supported by multiple cognitive computations and representations[27] across distributed and interconnected brain regions[28]. Variations in the type of aphasia reflect different patterns of damage to these brain regions and their interconnections. Therefore, recovery from aphasia must reflect changes to the remaining brain regions, their connections and the cognitive computations or representations that they support.

In classical conceptual models of language processing, different types of aphasia were assumed to reflect damage to input (posterior) areas (receptive aphasia)[2], output (anterior) areas (expressive aphasia)[1] or the connection between the two (conduction aphasia)[29]. As was initially suggested by Wernicke[30], contemporary models of language processing assume at least two major pathways (and potentially multiple subroutes)[31] within the dominant left hemisphere[32] (FIG. 1). The dorsal language pathway includes the connections between the superior temporal, inferior parietal and premotor regions, via the arcuate and superior longitudinal fasciculi. This pathway has been associated with the extraction and production of speech sound structures (that is, phonology, which is important for repetition and speech production)[32,33]. The ventral language pathway connects the middle and anterior temporal lobe to inferior frontal regions via multiple fasciculi. This pathway is crucial for translating to and from meaning, and is important for comprehension and semantically driven speech[32,33]. The 'dual pathways' model is supported by evidence from functional MRI (fMRI)[28], tractography[34], computational modelling[35] and lesion–symptom mapping. The last form of evidence demonstrates that the type of aphasia reflects the amount of damage to individual pathways; for example, conduction aphasia follows damage to the dorsal pathway and transcortical sensory aphasia follows damage to the ventral pathway[5,36].

Several aspects of healthy language processing that are relevant to recovery from aphasia are still not fully understood. First, it is unclear whether language-specific neurocomputations are performed unilaterally, bilaterally or asymmetrically, and whether the pathways in each hemisphere differ in their organization. Patient data imply that the left hemisphere is language dominant, as chronic aphasia is strongly associated with left hemisphere lesions[36]. However, the dorsal and ventral language comprehension pathways are bilateral, but left dominated, although the dorsal pathway is more strongly left-sided than the ventral pathway[33]. Accordingly, right hemisphere damage[37] or deactivation[38] can also result in language impairment. Furthermore, functional neuroimaging often reveals bilateral activation in healthy individuals during

Anomic aphasia
A type of mild acquired language deficit in which individuals have word-finding difficulties yet relatively preserved comprehension, repetition and speech production.

Non-invasive brain stimulation
(NIBS). A range of techniques, including transcranial magnetic stimulation and transcranial direct current stimulation, that can modulate activity in specific brain networks or regions using electromagnetic fields or electrical current.

Degeneracy
A term used to refer to brain regions or networks that are sufficient to perform a cognitive task but do so only when other structurally dissimilar networks that normally perform that task are damaged.

Variable neuro-displacement
The process whereby a neural network utilizes its spare capacity and increases its activity and/or performance in situations of increased difficulty. Under standard performance demands, activity in these areas is downregulated to save energy. Variable neuro-displacement aims to titrate performance against energy cost.

Domain-general, multidemand executive networks
Brain regions or networks that are activated across a variety of cognitive tasks or domains when task difficulty is increased.

Transcallosal disinhibition
The proposal that homologous regions in the two hemispheres try to inhibit each other's function and thus, following damage to one hemisphere, function in the contralateral region is released from this constraint.

comprehension tasks[39–41] and production tasks[42–44]. Lateralization of activation seems to vary between language activities, with propositional speech being more left hemisphere dominant than non-propositional speech, repetition, naming or comprehension, all of which are reliably associated with bilateral activation[45]. Understanding the basis of this variation in the lateralization of different language activities in healthy brain function will be foundational for explaining recovery from aphasia.

Second, virtually all right-handed individuals and more than 90% of left-handed individuals exhibit bilateral, albeit left-dominated, activation during connected speech production[46]. This finding suggests that handedness has a minimal role in determining hemispheric specialization for language in 99% of the population.

Third, although studies often refer to the right hemisphere language system as a single entity, it seems more likely that the right hemisphere contains several language pathways, similarly to the dominant left hemisphere, although the composition of these pathways is poorly understood.

Last, it is increasingly recognized that healthy and impaired language activities are supported not only by language-specific networks but also by non-language networks, including domain-general, multidemand executive networks[47–50] (FIG. 1a). Elucidating the role of these non-language networks in language processing will be crucial for understanding recovery from aphasia in terms of changing activity within these networks as well as the functional connectivity between them.

Box 1 | **Forms of neuroimaging data on recovery from aphasia**

Historically, information on the structure of the brain was obtained from post-mortem studies[1,2]. The arrival of in vivo neuroimaging, particularly MRI, radically improved the identification of grey and white matter damage[140]. Defining lesions in the early phase after stroke can be difficult, as in half of patients the infarct visible on T2-weighted MRI sequences is smaller in the subacute phase than in the chronic phase[141]. Diffusion tensor imaging can assess local white matter integrity[142] and long-range structural connectivity[143], both of which have been related to language and cognitive impairments after stroke[144–146].

Multiple methods are available to measure stroke-induced functional changes. Early techniques included single-photon emission CT[147] and PET[68], which have been superseded by functional MRI (fMRI), EEG and magnetoencephalography (MEG). Perfusion-weighted imaging techniques, such as arterial spin labelling, can be used to assess whether tissue is adequately perfused with oxygenated blood and to determine whether the level of perfusion relates to language recovery[148]. Importantly, although hypoperfusion directly compromises function, adequate perfusion does not imply that tissue is being used or is functioning optimally.

Repeated functional imaging can track function-related changes longitudinally[79,149] but is not without challenges. Neurovascular decoupling (breaking the normal link between neural activity and subsequent changes in cerebral blood flow), which reduces the blood oxygen level-dependent (BOLD) signal on fMRI, can occur in the acute phase, giving rise to a longitudinally changing BOLD signal that does not actually relate to the underlying neural activity[150]. This problem can be mitigated by controlling for cerebrovascular reactivity (for example, through breath-hold paradigms)[151] or by using MEG or EEG to measure electromagnetic neural signals directly[125]. In addition, activations might be epiphenomenal[109] or related to the increased cognitive effort involved in performing a task after neural damage[152]. Neural activation has a complex parabolic relationship with performance[108], such that poor and expert performance are both associated with lower activation (that is, failure to engage the system and neural efficiency[153,154], respectively), while moderate performance generates the greatest activation.

Recovery in the acute phase

Recovery from aphasia in the acute phase (the days immediately after stroke) could reflect the reversal of transient biological changes that temporarily dysregulate neuronal function[51]. Reperfusion of key language regions correlates with rapid resolution of aphasia in the first few days after stroke[52]. The complex inflammatory response to stroke, which suppresses neural function, peaks after approximately 1 week before resolving[53]. These examples of temporarily suppressed function can be simulated in computational models by transiently adjusting parameters to generate impaired performance, for example, by temporarily adding noise or reducing activation, and then reversing this adjustment. Indeed, this approach was adopted in a computational model of speech production, which was able to generate the differences between healthy and aphasic performance by adjusting parameters that controlled propagation of activation[54].

Recovery in subacute and chronic phases

Although acutely resolving biological changes clearly contribute to recovery in the first few days after stroke, they are not sufficient to account for recovery from aphasia during the subacute (first few weeks after onset) and chronic (starting at 3–6 months after onset) phases. During these later phases, changes to language performance are observed[16], for which no corresponding 'macroscopic' neurobiological changes, such as reperfusion or cellular regeneration, have been identified. We propose that existing theories of recovery from aphasia during the subacute and chronic phases can be conceptualized as specific examples of two key principles: degeneracy and variable neuro-displacement. These two principles are not mutually exclusive and might involve neural structures or networks in both hemispheres. In addition, the mechanisms of recovery considered in this section could begin during the acute phase.

For the various theories of recovery, we consider both the empirical evidence and any computationally implemented models. The latter are crucial because, as in other areas of clinical neuroscience[55], information from these models can increase the precision and rigour of our understanding of the neurocomputational bases of recovery. As we note in the following subsections, many prominent theories are currently limited to verbal redescriptions of observed phenomena and lack a computationally implemented mechanistic account. We conclude this section by critically discussing two issues that have had long-standing prominence in the literature — the role of the right hemisphere and the effect of transcallosal disinhibition during recovery from aphasia — in the context of the key principles of degeneracy and variable neuro-displacement.

Degeneracy

Degeneracy is a general notion that was applied first to complex biological systems and in recent times to the neural systems that underpin healthy brain function and neuropsychological impairments[56,57]. In basic terms, 'degeneracy' refers to any situation where the same outcome or the same function can arise from different elements or components. Everyday examples include the

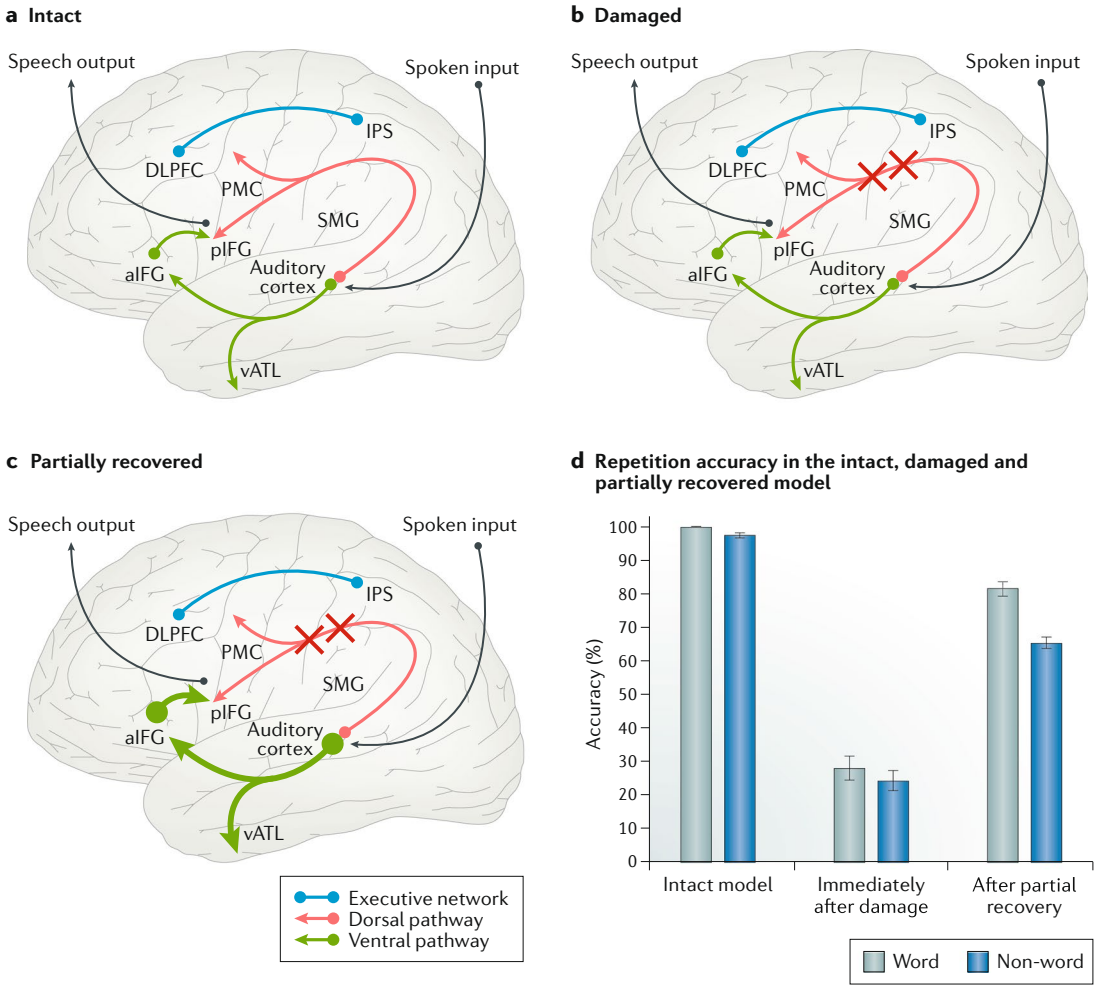


Fig. 1 | **The computationally implemented dual language pathways model. a** | In the dual language pathways model of language processing, the dorsal language pathway (red) connects the auditory cortex in the superior temporal region to the inferior parietal cortex (supramarginal gyrus (SMG)), premotor cortex (PMC) and posterior inferior frontal gyrus (pIFG), and is involved in phonological processing. The ventral language pathway (green) connects the auditory cortex to the ventral anterior temporal lobe (vATL) semantic hub and inferior frontal regions (anterior inferior frontal gyrus (aIFG)), and subserves semantic processing. The two language pathways interact with regions engaged by executively demanding tasks across multiple domains; this executive network (blue) involves regions of the dorsolateral prefrontal cortex (DLPFC) and intraparietal sulcus (IPS) in both hemispheres, as well as the anterior insula, dorsal anterior cingulate cortex and medial superior frontal cortex (not illustrated). Damage to different parts of the model generates the variety of aphasia subtypes. **b** | Following simulated damage to the dorsal language pathway (red crosses), the model exhibited conduction aphasia (poor repetition and speech production despite good comprehension). In this 'acute' phase immediately after simulated damage, the ventral language pathway remained unchanged. **c** | Partial recovery was simulated by re-exposing the damaged model to its learning environment. These simulated plasticity-related changes altered the division of labour across the pathways: in addition to retuning the remaining computational resources in the damaged dorsal pathway, the ventral language pathway became engaged by the repetition task (denoted by the broadened green arrow), which was not the case in the intact model. **d** | In the intact model, repetition of words (grey bars) and non-words (blue bars) was equally accurate. Immediately following damage to the dorsal pathway and before any recovery, word and non-word repetition showed equal impairment. Following partial recovery and the resultant change in the division of labour, repetition of words recovered more than repetition of non-words. Panel **d** adapted with permission from REF.[35], Elsevier.

fact that travelling between the same two locations can be achieved via different, non-overlapping routes or different modes of transport, or that the same meaning can be conveyed using very different words and sentence constructions. Degeneracy imbues resilience because when only a subset of elements are missing or compromised, the same goal can be achieved in a different way; for example, if your favourite road is blocked, you can use an alternative route to reach your destination. By extension, functional degeneracy in the context of cognition assumes the existence of multiple separate and structurally distinct neural networks, each of which is sufficient to perform a specific cognitive task[58]. In considering recovery from aphasia, we expand the definition of degeneracy to include neural networks that are not involved in the target cognitive task in healthy individuals but can be engaged for that task after damage, either immediately or following experience-dependent

plasticity (analogous to diverting cars along a bridleway when the main road has been blocked).

Although degeneracy is evident across multiple biological systems and everyday examples, its application to the link between brain and behaviour, and especially to the recovery of function following brain damage, needs to be specified in more detail, ideally with computational implementation to elucidate the mechanisms involved. These implementations will help us understand how distinct networks are formed and how they can generate the target behaviour, how the brain is able to dynamically combine and coordinate different networks to support the target function, what prompts the changes that occur after damage and how these changes might maximize the level of remaining function. In the subsections that follow, we discuss three different subtypes of degeneracy, which are not mutually exclusive.

Engagement of quiescent regions. Following damage to a network, degenerate regions that are quiescent in healthy individuals might become engaged to aid recovery. The hypotheses that quiescent regions of the right hemisphere become engaged in language processing during recovery from aphasia have existed since the nineteenth century[59]. These hypotheses suggest that right hemisphere homologues of left hemisphere language regions can perform similar — albeit weaker — language computations as their dominant left homologues and thus are engaged following left hemisphere damage[60]. Multiple neuroimaging studies have since uncovered right hemisphere activation in patients with post-stroke aphasia[61–63], and researchers have associated this activation with recovery of comprehension[64] and speech production[65–67]. For example, right inferior prefrontal regions are sometimes upregulated during partial recovery of propositional speech after damage to Broca's area (pars opercularis) in stroke[68] (FIG. 2a) and in individuals with low-grade glioma[69]. However, the mechanisms by which previously quiescent regions in either hemisphere could contribute to language function after stroke are not clear, and computational accounts of the process are lacking. For example, we do not know how quiescent areas learn the relevant representations and functions, how 'unused' representations in quiescent regions can remain dormant in healthy brains and are not repurposed for active use in other cognitive tasks and whether such quiescent regions are ever used in healthy function. How the overall system orchestrates the active and quiescent regions to generate healthy language function and to support recovery of language function in aphasia is also unclear.

Engagement of alternative language networks or pathways. The hypothesized engagement of alternative language networks during recovery from post-stroke aphasia is based on the notion that different processing pathways can be used to perform the same task and that recovery might reflect the use of remaining pathways. Engagement of alternative pathways could occur immediately after damage or following a period of computational adjustment. For example, in the dual pathways model of language[32,33], plasticity-related changes following damage to one pathway might change the relative contributions across pathways, resulting in the engagement of an undamaged pathway that was not previously involved in the function. This process was formally modelled in a computational implementation of the dual language pathways model[35]. The processing architecture of this model (FIG. 1a) was based on the known combination of dorsal language pathways, comprising frontoparietal regions and their connections, including the arcuate fasciculus, and ventral language pathways, comprising temporal and frontal lobe regions and their connections, including the middle longitudinal fasciculus and extreme capsule complex. Through repeated exposure to words and the application of an iterative learning algorithm to adjust the strengths of the model's connections, the model was able to learn to repeat and comprehend spoken words and generate speech from meaning.

Different types of aphasia were mimicked by simulating damage to different areas of the model, and plasticity-related recovery (via reoptimization of the remaining connection strengths) was induced by re-exposing the model to the words of the language and reapplying the learning algorithm. Following simulated damage to the dorsal language pathway (FIG. 1b), the model exhibited poor repetition but preserved comprehension, as is seen in patients with conduction aphasia. Following simulated plasticity-related partial recovery, in order to perform the repetition task the model not only used the remaining dorsal pathway resources but also engaged the ventral pathway, which the intact model did not use for this task (FIG. 1c). This engagement of the ventral pathway resulted in a 'side effect' in that repetition became better for real words (which have meaning) than non-words (which are meaningless)[35] (FIG. 1d) — a change that is also observed in patients with post-stroke aphasia[70]. Additional simulated damage to the ventral pathway in the partially recovered model (but not in the intact model) impaired repetition of real words, confirming the new-found reliance on the ventral pathway for word repetition[35].

Experimental evidence consistent with this type of degeneracy has been observed in healthy individuals following transcranial magnetic stimulation (TMS)-mediated inhibition of a key semantic region (the angular gyrus), which resulted in increased task-related activity in regions typically subserving phonology, including the supramarginal gyrus[71]. However, empirical evidence for this form of degeneracy in recovery from post-stroke aphasia (that is, observation of a neural network that is activated by a language task it was not involved in before damage) has not yet been collected. To complicate matters, distinct functional networks sometimes engage neuronal populations in spatially overlapping regions[72]. Future functional neuroimaging investigations of individuals with aphasia could explore multivariate analysis methods, including independent component analysis[73], that are capable of identifying distinct but spatially overlapping functional networks during language tasks.

Engagement of non-language regions. In its most extreme form, degeneracy might involve 'neurocomputational invasion', whereby a region that usually supports

Quiescent
Brain regions that are not activated during a language task in healthy individuals but can become activated and engaged by that language task after stroke are said to be quiescent.

Independent component analysis
A multivariate, data-driven analysis technique that can be used to decompose functional MRI data into statistically independent functional networks.

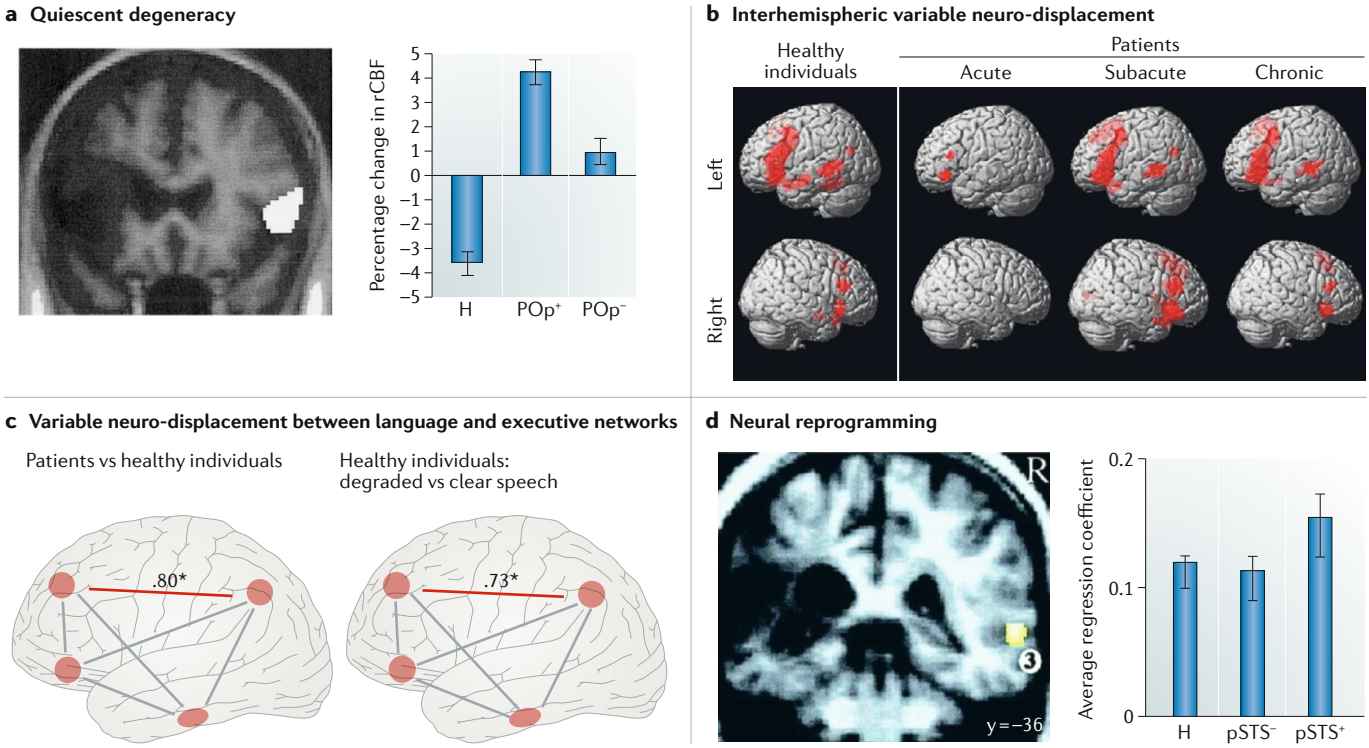


Fig. 2 | **Potential mechanisms of recovery from post-stroke aphasia.** Each panel shows a seminal piece of evidence for a specific mechanism of recovery. **a** | Quiescent degeneracy. Data showing that the right inferior frontal gyrus pars opercularis (POp), which is not involved (is quiescent) in speech production in healthy individuals, is sometimes used to support this language function following stroke. The left panel shows an area (white) in the right POp was activated more during a speech task in patients with damage to the left POp than in healthy individuals. The right panel shows that relative to a non-speech baseline, speech elicited significant deactivation in the right POp in healthy individuals (H) but significant activation in patients with left POp lesions (POp^+) and partial engagement in patients with spared left POp (POp^-). **b** | Interhemispheric variable neuro-displacement. Data that show a change in the division of labour across the hemispheres during the course of recovery from stroke. Clusters (red) of significant activation when listening to intelligible speech versus reversed speech in healthy individuals and in patients in the acute, subacute and chronic stages after left hemisphere stroke. Increased right hemisphere activation was observed in the subacute phase. **c** | Variable neuro-displacement between language and executive networks. Data from a key study that showed how non-language executive networks support language function in recovery from post-stroke aphasia as well as in healthy participants performing a challenging listening task. Functional connectivity between the superior frontal gyrus and the angular gyrus during speech comprehension was significantly greater in patients with left hemisphere stroke and aphasia than in healthy individuals (left panel; group Z score difference 0.80) and was significantly greater with degraded speech than with clear speech (right panel; group Z score difference 0.73). The asterisks indicate statistical significance. **d** | Neural reprogramming. An example of how recovery from post-stroke aphasia may reflect reprogramming of function within remaining language areas. After damage to the left posterior superior temporal sulcus (pSTS), the right hemisphere homologue shows increased physiological responsiveness during listening to speech (left panel; yellow, 3). The right panel shows the strength of association between increasing cerebral blood flow and word presentation rate in the right pSTS was greater in individuals with lesions affecting the left pSTS ($pSTS^+$) than in individuals with lesions sparing the left pSTS ($pSTS^-$) or healthy individuals (H). rCBF, regional cerebral blood flow. Panel **a** adapted with permission from REF.[68], Wiley © 2003 American Neurological Association. Panel **b** adapted with permission from Sharp et al.[104], Wiley © 2010 American Neurological Association. Panel **c** reprinted with permission from REF.[79], Oxford University Press. Panel **d** adapted with permission from Leff et al.[111], © 2002 Wiley-Liss, Inc.

a non-language activity is assimilated for language tasks following recovery from aphasia. Direct evidence for cortical plasticity has been observed in patients undergoing resection of glioma who, following repeated and extended direct electrocortical stimulation, exhibit relocation of motor and language regions[74]. Furthermore, cross-modal plasticity exists between auditory and visual cortices following complete sensory deprivation in healthy adults[75–78]. However, whether neurocomputational invasion occurs after stroke is unclear.

A seminal longitudinal investigation reported increased activity during auditory comprehension in the right inferior frontal gyrus (IFG) and right supplementary motor cortex in patients with mild aphasia 2 weeks after stroke as compared with healthy individuals (FIG. 2b). The magnitude of the change in activation of these areas between a few days and 2 weeks after stroke correlated positively with language recovery over the same period[79]. However, although activation of these right hemisphere regions was greater in the patients than in healthy individuals, activation occurred bilaterally in both groups; that is, no evidence of 'invasion' of novel regions was found. Furthermore, in the chronic phase, which is when neurocomputational invasion should theoretically be at its maximum, differences in right IFG activation were no longer detected between individuals with aphasia and healthy individuals[79]. To date, longitudinal studies investigating recovery from stroke-induced aphasia have included mainly individuals with mild aphasia, but to gain a full understanding of all types and severities

Pseudomodular
Modular cognitive systems comprise independent, fixed, discrete processing occurring in separate computational components. 'Pseudomodular' refers to a processing architecture that seems to be modular in form but can be reprogrammed to change functions within and between the computational components.

Spare capacity
The extent to which a neural network can increase its activity and/or performance in situations of increased task difficulty.

Distributed representations
Information coded across multiple processing units within computational models or across multiple areas of the brain.

Graceful degradation
A nonlinear pattern of decline in which performance is minimally reduced at low to moderate levels of damage.

of aphasia, future investigations will need to include individuals with moderate and severe aphasia.

One computational model has explored how the healthy language system could form pseudomodular networks during development, which can be adjusted through learning and plasticity after damage[80] (BOX 2). A key observation from this model is that assimilation of non-language regions does not necessarily reflect regions completely switching function but rather signifies 'interdigitated' infiltration of language computations alongside the existing function in the 'invaded' cortical region.

Variable neuro-displacement

The principle of degeneracy is borrowed from biology[56], but we can also look to other disciplines for important mechanistic accounts of systems that are under pressure or have been damaged. For example, a core tenet of engineering is to design a system that is resilient to functional stresses. Engineering is also concerned with the balance between performance and cost; this concept certainly applies to the brain, which is an extremely metabolically expensive organ[81]. Thus, it seems likely that neurocognitive systems should have evolved to be tolerant to variable levels of performance demand, and to titrate performance against metabolic energy demands[82]. Accordingly, under standard levels of performance demand, the full language system should be downregulated to save energy but have spare capacity that can be utilized when necessary; this process is known as 'variable displacement' in modern engine design[83].

Variable neuro-displacement might plausibly co-occur with degeneracy during recovery from aphasia and, as such, these two principles should not be considered mutually exclusive (BOX 3). However, variable neuro-displacement does provide an alternative explanation for apparently quiescent brain regions which, in fact, might be intrinsic to healthy language networks but downregulated unless the system is under pressure (for example, following damage, or when a healthy individual performs a difficult task), rather than being additional areas that are co-opted de novo as a form of degeneracy. Therefore, one source of recovery from aphasia might be the inherent neurocomputational resilience of the well-engineered cognitive system. Indeed, such resilience might be analogous to cognitive reserve in healthy ageing and dementia (BOX 4). The subsections that follow outline the ways in which spare capacity could exist within or between neural networks.

Variable neuro-displacement within unaffected parts of the damaged system. Neural systems frequently have spare capacity; for example, in Parkinson disease, 50% of the dopaminergic neurons in the substantia nigra can be lost before motor symptoms develop[84]. Likewise, a computational model of the semantic system that incorporated spare capacity, distributed representations and training with noise exhibited graceful degradation in response to damage[85]. Spare capacity within neural systems provides immediate resilience (without the need for retraining), but also provides the opportunity for plasticity-related recovery within the remaining parts of the damaged network[35].

Following stroke-induced damage, upregulated activation in the perilesional area could reflect the kind of variable neuro-displacement described above, if we assume that the spared tissue was part of the functional neural network before damage. In keeping with this hypothesis, increased perilesional activity has been associated with improved long-term language outcomes after stroke[44,79,86–90]. For instance, in one study, increased activation in left parietal and premotor regions was associated with improved naming following a cueing-based therapy in the chronic phase[91]. However, the studies to date that have associated left hemisphere activation with language outcomes have not controlled for lesion size; therefore, the results might have been confounded by individuals with larger lesions who had less perilesional tissue to activate.

Given that single cognitive functions tend to involve distributed regions in both hemispheres[40], the areas beyond the immediate vicinity of the lesion might also be important in recovery from aphasia. Evidence for spare capacity in the distributed network that underpins language function in healthy individuals has come from both neuropsychology and brain stimulation research. For example, individuals with left temporal lobe resections showed increased activation of the right anterior temporal lobe during a semantic decision task[92]. Similarly, in healthy participants, inhibition of the left anterior temporal lobe by TMS upregulated activity in the right anterior temporal lobe[93,94], and the degree of upregulation predicted the reaction time during semantic tasks. Importantly, this upregulation is unlikely to reflect degeneracy via the recruitment of quiescent regions or neurocomputational invasion because the same result was obtained when healthy participants completed a more demanding version of the same task[92]. Instead, it indicates that spare capacity makes the bilateral semantic network inherently adaptive and thus resilient to changing demands or unilateral damage.

Box 2 | **A model of neurocomputational invasion**

A computational model explored how neurocomputational invasion into non-language regions might be achieved. The model consisted of two parallel networks that were simultaneously trained on one of two independent tasks[80]. Sparse connectivity between networks resulted in modularity, with each network becoming specialized to perform one of the two tasks. Sudden versus gradual damage to one of the two networks, followed by extensive retraining (to reflect experience-dependent plasticity), was used to simulate the effect of fast (stroke-like) versus slow (tumour-like) damage on subsequent recovery. Analyses of the model showed that this simulated recovery represented degeneracy through interdigitated quiescent capacity: the relatively underused nodes in the undamaged network were adopted to take over the network's impaired function, whereas the highly used nodes remained tied to the undamaged non-language function[80]. Thus, neurocomputational invasion into the undamaged network was possible only because the weak connections between the damaged and undamaged networks were quiescent before damage but allowed plasticity after damage[80]. This finding is in keeping with neurosurgical mapping evidence that neuroplasticity is highest in the cortical grey matter regions that have significant connectivity to other cortical regions[155]. For maximal neurocomputational invasion to occur, the time course of damage and thus plasticity needs to be gradual rather than sudden. This finding is consistent with the experimental observation that slow-growing (but not fast-growing) left-sided gliomas result in the involvement of the right inferior frontal gyrus in speech production[156].

Box 3 | **Degeneracy or variable neuro-displacement?**

We propose that variable neuro-displacement and degeneracy might co-occur during recovery from aphasia, raising the question of how we can distinguish between these mechanisms. Definitive answers to this question remain elusive owing to the difficulty of making specific predictions to be tested with patients' neuroimaging and behavioural data without the use of implemented computational models. The principle of degeneracy assumes the existence of brain regions that are rarely, if ever, used in healthy individuals but can perform computations relevant to a language task after stroke, either immediately or following experience-dependent plasticity.

Variable neuro-displacement assumes that networks involved in the healthy language system have spare capacity, and the same intrinsic mechanisms that can utilize this spare capacity in demanding situations for healthy individuals also support recovered function in patients. This principle implies that the changes in brain activation and functional connectivity observed in patients with stroke-induced aphasia should be seen in healthy individuals if a task is sufficiently challenging. If variable neuro-displacement is the only mechanism underpinning recovery from aphasia, then a sufficient explanation of the patient data should be derivable from a complete exploration of the intrinsic resilience and capacity of the distributed networks involved in the language system in healthy individuals.

Variable neuro-displacement across language networks. In healthy individuals, networks within the language system are at equilibrium, balancing their computational resources against performance demands. However, when the system is partially damaged, resources are reduced and the remaining networks are not immediately configured to provide the optimal solution for balancing the reduced resources against performance demands. Accordingly, recovery can be viewed as a reoptimization process that leads to a new equilibrium, and can involve changes in the division of labour across language networks that are comparable to the engagement of alternative networks already noted in the section entitled 'Degeneracy'.

Two types of variable neuro-displacement across networks can be envisaged. In one scenario, a network is truly quiescent in the premorbid state and begins to contribute only after damage. In the second scenario, multiple networks make differential contributions to a core language activity in the premorbid state and the division of labour is adjusted after damage to support recovery. For example, the triangle computational model of reading aloud has both a direct pathway that maps orthography to phonology and an indirect pathway via semantics[27]. During recovery after damage in this model, the division of labour was found to shift between the direct and semantically mediated reading pathways[95]. The pattern of deficits observed after plasticity-related recovery of the triangle model was consistent with that observed in patients with central alexias[95]. Similarly, evidence from EEG recordings indicates that individuals with agrammatic aphasia exhibit more meaning-related waveforms than individuals with normal grammatical function. This finding suggests that individuals with agrammatic aphasia attempt to compensate for impaired grammar by increasing their reliance on semantics[96].

Triangle computational model of reading aloud
An implemented computational model of reading aloud that includes three interconnected representational systems: orthography (written word forms), phonology (the sound structure of words) and semantics (word meaning).

Variable neuro-displacement between language and non-language networks. Evidence that language regions interact with domain-general executive-demand networks is accumulating[97] (FIG. 1a). In healthy individuals, demanding tasks require maximal engagement of the executive-demand network[98], and following focal damage to frontal regions of this network, activity is upregulated in undamaged parietal regions[99]. Therefore, variable neuro-displacement between language and executive-control cognitive systems might occur during recovery from aphasia. In keeping with this hypothesis, performance in behavioural tasks that assess top-down control processes, such as attention[100,101], processing speed[102] and cognitive flexibility[103], frequently correlates with language ability after stroke.

Evidence from neuroimaging studies also indicates a greater involvement of domain-general regions in language function in patients with stroke than in healthy individuals. For example, increased connectivity was detected between the superior frontal gyrus (part of the frontoparietal executive network) and the angular gyrus during a comprehension task in patients with chronic post-stroke aphasia but not in healthy individuals performing the same task (FIG. 2c), and the degree of frontoparietal integration correlated positively with language recovery[104]. Likewise, several subsequent studies found increased activation in midline frontal regions during various language activities in patients with chronic or recovering post-stroke aphasia but not in healthy individuals[49,73,105]. Importantly, when task difficulty was increased in healthy participants, activation changes were observed that were similar to those seen in individuals with aphasia[49,104]. Therefore, these increases in activation are likely to be a result of variable neuro-displacement and the use of spare capacity rather than degeneracy via recruitment of quiescent regions or neurocomputational invasion (FIG. 2c). To fully understand this form of variable displacement, computational models need to be generated to provide mechanistic accounts of how language-specific and domain-general networks interact, both in healthy performance and during recovery from aphasia.

Right hemisphere theories

Theories regarding the role of the right hemisphere in recovery from post-stroke aphasia have existed since the nineteenth century[59], when Barlow reported the case of an individual who recovered from and then redeveloped aphasia and was found to have bilateral IFG lesions after death[106]. This finding was interpreted by Gowers as meaning that the right IFG had taken over the function lost after damage to the left Broca's area, but this recovered function was undermined by the subsequent right hemisphere stroke[107]. As stated earlier, we consider right hemisphere 'theories' to be a subset of verbally described hypotheses about the regions that might be associated with language recovery. Typically, these hypotheses are not described in relation to a computational or mechanistic account, but they could represent a specific example of the degeneracy or variable neuro-displacement principles outlined earlier rather than being separate recovery mechanisms in themselves. The current literature contains opposing opinions about the utility of right hemisphere involvement: some researchers consider it to be positive[62,64], whereas others have proposed that right hemisphere involvement might be maladaptive

Parametric correlation
An approach used in some functional neuroimaging studies which involves varying the parameter of interest (for example, speech rate) in a graded way and exploring which brain regions show activity changes that correlate with that parameter.

Multivoxel pattern analysis
A multivariate analysis technique that takes into account spatial patterns of activity across multiple brain voxels rather than assuming activity in each voxel is independent.

Representational similarity analysis
A multivariate analysis technique that calculates similarities between multivoxel functional MRI responses to different stimulus representations.

Dynamic causal modelling
A method of analysing functional neuroimaging data that infers causal interactions between brain regions (effective connectivity) rather than looking only for statistical correlations between their activity (functional connectivity).

(for example, the transcallosal disinhibition hypothesis described below). Both beliefs have been used to guide the development of therapeutic interventions.

Multiple functional neuroimaging studies have observed right hemisphere activation in patients with post-stroke aphasia[61–63]. However, changes in activation are difficult to interpret as the magnitude of activation is unlikely to have a linear relationship with functional performance[108], and activation alone does not prove a contribution of the activated region to performance[109]. A handful of studies have used TMS to investigate the causal nature of right hemisphere activation after stroke. One such study showed that transient inhibition of the activated right hemisphere regions impaired semantic performance[110]. Advances in fMRI analytics have allowed the investigation of changes in haemodynamic response function. For instance, by moving away from simple subtraction analyses towards parametric correlation, one seminal study showed that following left posterior superior temporal sulcus lesioning, the response curve between the rate of speech presentation and neural activity became steeper in the right posterior superior temporal sulcus[111] (FIG. 2d). This increased gradient implies that recovery of function can follow from reprogramming brain function within existing regions rather than necessitating the engagement of additional areas. Crucially, this change in physiological responsiveness would not have been apparent in univariate contrasts of fMRI activation alone[111].

Other advances in imaging analysis, including connectivity analyses[112], multivoxel pattern analysis[113] and representational similarity analysis[114], are likely to provide novel insights regarding the information conveyed in right hemisphere (or other newly activated) areas after recovery from aphasia. Indeed, a seminal study using multivoxel pattern analysis found that the pattern of activation in the right frontal cortex during auditory comprehension was predictive of subsequent language recovery[115], although this finding in itself does not reveal the nature of the computations being performed by this region or how they contribute to recovery.

In some studies, activation in the right hemisphere has been associated with poorer language performance in the chronic phase after stroke[86,88,116]. Negative associations between right hemisphere activation and language performance are sometimes considered to support a 'regional hierarchy' theory of recovery[117] in which optimal language recovery is associated with left-dominant activation, whereas engagement of right hemisphere networks is a suboptimal second option. Evidence from the motor cortex suggests that transcallosal inhibitory connections can be 'released' (that is, disinhibited) after TMS[118]. This finding has been used to support the hypothesis that transcallosal disinhibition might also occur in hemispatial neglect[119] and aphasia[117,120]. For example, right hemisphere upregulation in post-stroke aphasia has been proposed to result from disinhibition of redundant secondary language networks that are always present in the right hemisphere but in healthy individuals are kept inactive by transcallosal inhibition from the dominant hemisphere[117]. In turn, upregulated right hemisphere language networks transcallosally might inhibit the spared left hemisphere homologues that are needed for optimal language performance[117]. Therefore, these 'juvenile' right hemisphere systems might need to be suppressed to allow function to return to the left hemisphere. Motivated by this general hypothesis, one study found that the response to speech therapy was augmented when TMS was administered to the right inferior frontal region[120].

Negative associations between right hemisphere activation and language might partly reflect the relationship between the severity of aphasia and lesion size[65], as smaller lesions tend to cause milder aphasia and also leave more of the left hemisphere intact and able to be activated. Moreover, there is considerable evidence that the right hemisphere is activated during healthy language performance[43,94], and that improved recovery of comprehension[64,111] or speech production[65–67] ability after stroke is associated with augmented or reprogrammed[111] right hemisphere activation. Indeed, one study of the semantic system in healthy individuals found evidence of positive effective connectivity (that is, an inferred causal connection between regions) between the left and right anterior temporal lobes, which was upregulated following left anterior temporal lobe TMS, perhaps reflecting variable neuro-displacement[94].

Although transcallosal disinhibition is a commonly repeated hypothesis, we are not aware of any evidence that transcallosal inhibitory wiring exists outside the primary motor system in which it was first described[118], nor are we aware of any evidence from dynamic causal modelling that demonstrates increased negative effective connectivity from contralesional to ipsilesional cortical regions during language or attention tasks after stroke. The notion of transcallosal inhibition is puzzling from a computational perspective, as how right hemisphere systems can develop in the presence of persistent inhibition is unclear. Secondary right hemisphere networks that are seldom if ever used would seem to constitute a biologically expensive back-up system[81].

Box 4 | **Resilience and cognitive reserve**

'Cognitive reserve' is a term that refers to the resilience of a system to damage[157] and is often used in the context of healthy ageing and dementia to describe some form of spare capacity or compensatory system that mitigates the effects of brain dysfunction or mild brain damage. The mechanisms underlying cognitive reserve are unclear, although they could be related to variable neuro-displacement. Typically, cognitive reserve has been associated with higher premorbid IQ and/or better multidemand executive ability[158]. This association has similarities to notions of the importance of non-language executive networks in recovery of language after stroke. Cognitive reserve could also reflect the positive impact of greater computational resources (for example, more spare capacity) within the damaged networks themselves. Indeed, additional pathology such as cerebral small vessel disease[159] or atrophy of undamaged cortex might impede recovery after stroke by reducing spare capacity[160].

Most computational accounts of recovery are based around an 'average' system or profile, but individuals show important variations in premorbid abilities, connectivity and neurocomputational set-up[161]. These variations are likely to lead to interindividual differences both in the overall potential for recovery and the mechanisms involved. Interindividual differences are directly relevant to the accounts of degeneracy; that is, there could be variations in the range and type of brain regions, networks and computational processes that are available to perform the same task. These differences also have implications for the variable neuro-displacement hypothesis; for example, differences in the division of labour that is found across networks or in the resources available, which would change the spare resources in those networks and thus their capacity for plasticity-related improvements.

Melodic intonation therapy
A type of speech and language therapy that uses music to encourage fluent speech production through improved intonation and rhythm.

Phonotactic statistics
The pattern and frequency of the sound sequences that are found in a language.

More broadly, computational accounts of the function of the right hemisphere in recovery from aphasia are lacking. Future studies will need to establish how the networks in the left and right hemispheres work together in healthy individuals, how the functions and representations in these hemispheres develop, how bilateral networks might be beneficial and whether good recovery reflects reversion to the damaged left hemisphere or upregulation of right hemisphere networks. One initial computational exploration of a bilateral semantic network found that bilaterality makes a system more robust to unilateral perturbation or damage, and enables the system to recover function through plasticity-based learning[121].

Implications for treatment

A clearer understanding of the mechanisms underlying normal and recovered language performance should facilitate the design of therapeutic approaches to ensure maximal spontaneous recovery of language function and efficacy of interventions after stroke. In this section we discuss some key implications of our current understanding of the mechanisms of recovery from aphasia for the application of pharmacological, SLT and NIBS interventions.

Pharmacological interventions

By increasing our knowledge of the neurocomputational mechanisms involved in spontaneous recovery, we might be able to develop pharmacological interventions to modify the appropriate neurotransmitter systems. Upregulation of the monoaminergic pathways that are thought to influence attention and working memory might be beneficial if variable neuro-displacement in domain-general executive networks is crucial for recovery from aphasia[122]. If connection reweighting to shift the division of labour between the dorsal and ventral language pathways or to allow neurocomputational invasion into new cortical regions is crucial for recovery, then pharmacologically increasing experience-dependent plasticity might improve outcomes.

Selective serotonin reuptake inhibitors have been associated with language recovery above and beyond any effects on mood in patients with subacute post-stroke aphasia[25], and may increase neuronal plasticity[123]. Cholinergic input to the cortex from the nucleus basalis is important for plasticity associated with complex cognitive processing[124]; however, a trial of an acetylcholinesterase inhibitor found a detrimental effect of the drug on comprehension after stroke[125]. The results of this trial highlight the need for further studies to investigate both positive and negative effects of pharmacological interventions. An important consideration is how best to use antiepileptic drugs for the treatment of seizures after stroke, as these drugs decrease cortical excitability, and the antiepileptic phenytoin is well known to worsen cognitive outcomes after certain forms of brain injury[126].

Speech and language therapy

SLT uses many different kinds of intervention, including behaviourally focused neurorehabilitation techniques, which are designed to identify and target the specific aspects of language that have been impaired after brain damage, or to encourage compensatory mechanisms and strategies. If executive or attention systems are important in variable neuro-displacement and thus recovery, we might need 'motivating' SLT to ensure a suitably engaging learning environment. For example, the entertaining and game-like nature — or 'gamification' — of recently developed therapies and apps for mobile devices might make them more engaging or demanding than plain repetitive practice approaches, thereby helping to increase executive system involvement in language function[125,127,128]. If recovery occurs by shifting the division of labour between different language networks[95], SLT could be targeted at the undamaged language pathway; for example, engagement of the ventral language pathway for repetition might be increased by concurrent priming or retrieval of word meanings. Establishing a true mechanistic understanding of the role of the right hemisphere in language function after stroke will have significant implications for the use of some current and future forms of SLT; for example, melodic intonation therapy is thought to work by engaging right hemisphere brain regions[129].

Some computational models indicate that plasticity in general is maximal at the beginning of recovery and decreases as learning occurs[95], and early, high-dosage SLT might be needed during this 'critical period'. However, controversy remains over whether intensive therapy administered for a limited period[130] or distributed bursts of therapy that are spaced out over time are optimal during the chronic phase of recovery, when plasticity is reduced[95]. Evidence suggests that when the overall dose of therapy is the same, a distributed pattern is better than intensive therapy (typically, intensive therapies tend to include a higher dosage than distributed interventions)[131], which is in keeping with psychological models of distributed learning[132]. As a final example, computational models of reading suggest that if the training set is not diverse enough, substandard sampling of phonotactic statistics occurs and the system is unable to generalize its learning to read non-words[133]. This finding suggests that we need to ensure that the learning environment used for SLT resamples diverse language representations to allow optimal recovery.

Non-invasive brain stimulation

To know which regions and/or networks to target with NIBS, we need a better understanding of the mechanisms underlying recovery from aphasia. Many NIBS studies have attempted to aid recovery from aphasia by inhibiting the right IFG pars triangularis with the aim of disinhibiting the left IFG, an approach that is based on theories of transcallosal disinhibition. Evidence is growing that such NIBS protocols can aid the recovery of naming ability after stroke[134,135]. One trial reported increased left hemisphere activation following low-frequency, inhibitory repetitive TMS of the right IFG pars triangularis compared with sham stimulation[120]. However, as detailed previously, we are not aware of convincing evidence for the existence of transcallosal inhibitory wiring between language areas in healthy individuals. Furthermore, the same TMS protocol can induce highly variable responses in different individuals[136,137]. Therefore, it is of paramount importance that

future research elucidates whether inhibitory TMS protocols applied to the right IFG actually inhibit activation of the right IFG and disinhibit the left IFG, whether any such changes are due to reduced negative effective connectivity from the right IFG to the left IFG and whether the magnitude of such changes is associated with the degree of language recovery.

If executive and/or attention systems are important in recovery, we might need to target domain-general networks with NIBS; indeed, TMS of midline frontal regions has been found to improve performance during artificial vocabulary learning in healthy individuals[138], although this effect has not yet been investigated in the context of recovery from aphasia. If variable neuro-displacement into the spare capacity of language networks is important for recovery from aphasia, then stimulation of perilesional left hemisphere regions might aid recovery; however, current evidence for the efficacy of this method is limited[139]. More broadly, if we are to interpret clinical responses to NIBS observed at the group level, we must improve our understanding of NIBS methods and their impact on activity in local and distal neural networks.

Conclusions

Despite the increasing number of studies investigating the neural basis of recovery from aphasia after stroke, many theories are still little more than verbal descriptions of observed phenomena. Computational models suggest that the mechanisms underlying recovery from post-stroke aphasia can be conceptualized as the engagement of degenerate neural networks or the use of spare capacity within or between networks via variable neuro-displacement. Variable neuro-displacement might involve language-specific systems or non-language, domain-general cognitive systems that are involved in executive function. This conceptual framework is relevant to the design of pharmacological, SLT and NIBS interventions that aim to maximize language recovery after stroke. Longitudinal neuroimaging data on recovery from aphasia as well as different (notably multivariate) types of imaging analysis are needed to better elucidate the computations performed by activated cortical regions after stroke.

Published online: 26 November 2019

1. Broca, P. Sur le siège de la faculté du language articulé. *Bull Soc. Anthropol.* **6**, 337–393 (1865).
2. Wernicke, C. *Der aphasische Symptomencomplex, eine psychologische Studie auf anatomischer Basis* (Cohn and Weigert, 1874).
3. Butler, R. A., Lambon Ralph, M. A. & Woollams, A. M. Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. *Brain* **137**, 3248–3266 (2014).
4. Lacey, E. H., Skipper-Kallal, L. M., Xing, S., Fama, M. E. & Turkeltaub, P. E. Mapping common aphasia assessments to underlying cognitive processes and their neural substrates. *Neurorehabil. Neural Repair* **31**, 442–450 (2017).
5. Mirman, D. et al. Neural organization of spoken language revealed by lesion-symptom mapping. *Nat. Commun.* **6**, 6762 (2015).
6. Fridriksson, J. et al. Revealing the dual streams of speech processing. *Proc. Natl Acad. Sci. USA* **113**, 15108–15113 (2016).
7. Benjamin, E. J. et al. Heart disease and stroke statistics—2017 update: a report from the American Heart Association. *Circulation* **135**, e146–e603 (2017).
8. Engelter, S. T. et al. Epidemiology of aphasia attributable to first ischaemic stroke: incidence, severity, fluency, etiology, and thrombolysis. *Stroke* **37**, 1379–1384 (2006).
9. Boehme, A. K., Martin-Schild, S., Marshall, R. S. & Lazar, R. M. Effect of aphasia on acute stroke outcomes. *Neurology* **87**, 2348–2354 (2016).
10. Ellis, C., Simpson, A. N., Bonilha, H., Mauldin, P. D. & Simpson, K. N. The one-year attributable cost of poststroke aphasia. *Stroke* **43**, 1429–1431 (2012).
11. Tsouli, S., Kyritsis, A. P., Tsagalis, G., Virvidaki, E. & Vemmos, K. N. Significance of aphasia after first-ever acute stroke: impact on early and late outcomes. *Neuroepidemiology* **33**, 96–102 (2009).
12. Lomas, J. & Kertesz, A. Patterns of spontaneous recovery in aphasic groups: a study of adult stroke patients. *Brain Lang.* **5**, 388–401 (1978).
13. Yagata, S. A. et al. Rapid recovery from aphasia after infarction of Wernicke's area. *Aphasiology* **31**, 951–980 (2017).
14. Maas, M. B. et al. The prognosis for aphasia in stroke. *J. Stroke Cerebrovasc. Dis.* **21**, 350–357 (2012).
15. Pedersen, P. M., Jørgensen, H. S., Nakayama, H., Raaschou, H. O. & Olsen, T. S. Aphasia in acute stroke: incidence, determinants, and recovery. *Ann. Neurol.* **38**, 659–666 (1995).
16. Hope, T. M. H. et al. Right hemisphere structural adaptation and changing language skills years after left hemisphere stroke. *Brain* **140**, 1718–1728 (2017).
17. Elkana, O., Frost, R., Kramer, U., Ben-Bashat, D. & Schweiger, A. Cerebral language reorganization in the chronic stage of recovery: a longitudinal fMRI study. *Cortex* **49**, 71–81 (2013).
18. Laska, A. C., Hellblom, A., Murray, V., Kahan, T. & Von Arbin, M. Aphasia in acute stroke and relation to outcome. *J. Intern. Med.* **249**, 413–422 (2001).
19. El Hachioui, H. et al. Screening tests for aphasia in patients with stroke: a systematic review. *J. Neurol.* **264**, 211–220 (2017).
20. Wade, D. T., Hewer, R. L., David, R. M. & Enderby, P. M. Aphasia after stroke: natural history and associated deficits. *J. Neurol. Neurosurg. Psychiatry* **49**, 11–16 (1986).
21. Gilmore, N., Meier, E. L., Johnson, J. P. & Kiran, S. Non-linguistic cognitive factors predict treatment-induced recovery in chronic post-stroke aphasia. *Arch. Phys. Med. Rehabil.* **100**, 1251–1258 (2019).
22. Seghier, M. L. et al. The PLORAS database: a data repository for predicting language outcome and recovery after stroke. *Neuroimage* **124**, 1208–1212 (2016).
23. Ramsey, L. E. et al. Behavioural clusters and predictors of performance during recovery from stroke. *Nat. Hum. Behav.* **1**, 38 (2017).
24. Siegel, J. S. et al. Re-emergence of modular brain networks in stroke recovery. *Cortex* **101**, 44–59 (2018).
25. Hillis, A. E. et al. Predicting recovery in acute poststroke aphasia. *Ann. Neurol.* **83**, 612–622 (2018).
26. Goodglass, H. & Kaplan, E. in *The Assessment of Aphasia and Related Disorders* (Lea & Febiger, 1983).
27. Patterson, K. & Ralph, M. A. Selective disorders of reading? *Curr. Opin. Neurobiol.* **9**, 235–239 (1999).
28. Saur, D. et al. Ventral and dorsal pathways for language. *Proc. Natl Acad. Sci. USA* **105**, 18035–18040 (2008).
29. Heilman, K. M. Aphasia and the diagram makers revisited: an update of information processing models. *J. Clin. Neurol.* **2**, 149–162 (2006).
30. Weiller, C., Bormann, T., Saur, D., Musso, M. & Rijntjes, M. How the ventral pathway got lost: and what its recovery might mean. *Brain Lang.* **118**, 29–39 (2011).
31. Friederici, A. D. & Gierhan, S. M. The language network. *Curr. Opin. Neurobiol.* **23**, 250–254 (2013).
32. Hickok, G. & Poeppel, D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* **92**, 67–99 (2004).
33. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).
34. Saur, D. et al. Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *Neuroimage* **49**, 3187–3197 (2010).
35. Ueno, T., Saito, S., Rogers, T. & Lambon Ralph, M. Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron* **72**, 385–396 (2011).
36. Kummerer, D. et al. Damage to ventral and dorsal language pathways in acute aphasia. *Brain* **136**, 619–629 (2013).
37. Gajardo-Vidal, A. et al. How right hemisphere damage after stroke can impair speech comprehension. *Brain* **141**, 3389–3404 (2018).
38. Hickok, G. et al. Bilateral capacity for speech sound processing in auditory comprehension: evidence from Wada procedures. *Brain Lang.* **107**, 179–184 (2008).
39. Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D. & Wise, R. J. Temporal lobe regions engaged during normal speech comprehension. *Brain* **126**, 1193–1201 (2003).
40. Rice, G. E., Lambon Ralph, M. A. & Hoffman, P. The roles of left versus right anterior temporal lobes in conceptual knowledge: an ALE meta-analysis of 97 functional neuroimaging studies. *Cereb. Cortex* **25**, 4374–4391 (2015).
41. Halai, A. D., Parkes, L. M. & Welbourne, S. R. Dual-echo fMRI can detect activations in inferior temporal lobe during intelligible speech comprehension. *Neuroimage* **122**, 214–221 (2015).
42. Fridriksson, J. et al. Modulation of frontal lobe speech areas associated with the production and perception of speech movements. *J. Speech Lang. Hear. Res.* **52**, 812–819 (2009).
43. Fedorenko, E., Behr, M. K. & Kanwisher, N. Functional specificity for high-level linguistic processing in the human brain. *Proc. Natl Acad. Sci. USA* **108**, 16428–16433 (2011).
44. Warburton, E., Price, C. J., Swinburn, K. & Wise, R. J. Mechanisms of recovery from aphasia: evidence from positron emission tomography studies. *J. Neurol. Neurosurg. Psychiatry* **66**, 155–161 (1999).
45. Walenski, M., Europa, E., Caplan, D. & Thompson, C. K. Neural networks for sentence comprehension and production: an ALE-based meta-analysis of neuroimaging studies. *Hum. Brain Mapp.* **40**, 2275–2304 (2019).
46. Mazoyer, B. et al. Gaussian mixture modelling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PLOS ONE* **9**, e101165 (2014).
47. Gordon, P. C., Hendrick, R. & Levine, W. H. Memory-load interference in syntactic processing. *Psychol. Sci.* **13**, 425–430 (2002).
48. Carretti, B., Borella, E., Cornoldi, C. & De Beni, R. Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: a meta-analysis. *Learn. Individ. Differ.* **19**, 246–251 (2009).

49. Brownsett, S. L. et al. Cognitive control and its impact on recovery from aphasic stroke. *Brain* **137**, 242–254 (2014).
50. Mitchell, R. L., Vidaki, K. & Lavidor, M. The role of left and right dorsolateral prefrontal cortex in semantic processing: a transcranial direct current stimulation study. *Neuropsychologia* **91**, 480–489 (2016).
51. Szalay, G. et al. Microglia protect against brain injury and their selective elimination dysregulates neuronal network activity after stroke. *Nat. Commun.* **7**, 11499 (2016).
52. Hillis, A. E. & Heidler, J. Mechanisms of early aphasia recovery. *Aphasiology* **16**, 885–895 (2002).
53. Fu, Y., Liu, Q., Anrather, J. & Shi, F. D. Immune interventions in stroke. *Nat. Rev. Neurol.* **11**, 524–535 (2015).
54. Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M. & Gagnon, D. A. Lexical access in aphasic and nonaphasic speakers. *Psychol. Rev.* **104**, 801–838 (1997).
55. Teufel, C. & Fletcher, P. C. The promises and pitfalls of applying computational models to neurological and psychiatric disorders. *Brain* **139**, 2600–2608 (2016).
56. Tononi, G., Sporns, O. & Edelman, G. M. Measures of degeneracy and redundancy in biological networks. *Proc. Natl Acad. Sci. USA* **96**, 3257–3262 (1999).
57. Edelman, G. M. & Gally, J. A. Degeneracy and complexity in biological systems. *Proc. Natl Acad. Sci. USA* **98**, 13763–13768 (2001).
58. Price, C. J. & Friston, K. J. Degeneracy and cognitive anatomy. *Trends Cogn. Sci.* **6**, 416–421 (2002).
59. Finger, S., Buckner, R. L. & Buckingham, H. Does the right hemisphere take over after damage to Broca's area? the Barlow case of 1877 and its history. *Brain Lang.* **85**, 385–395 (2003).
60. Grafman, J. Conceptualizing functional neuroplasticity. *J. Commun. Disord.* **33**, 345–356 (2000).
61. Qiu, W. H. et al. Evidence of cortical reorganization of language networks after stroke with subacute Broca's aphasia: a blood oxygenation level dependent-functional magnetic resonance imaging study. *Neural. Regen. Res.* **12**, 109–117 (2017).
62. Robson, H. et al. The anterior temporal lobes support residual comprehension in Wernicke's aphasia. *Brain* **137**, 931–943 (2014).
63. Turkeltaub, P. E., Messing, S., Norise, C. & Hamilton, R. H. Are networks for residual language function and recovery consistent across aphasic patients? *Neurology* **76**, 1726–1734 (2011).
64. Crinion, J. & Price, C. J. Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. *Brain* **128**, 2858–2871 (2005).
65. Griffis, J. C. et al. The canonical semantic network supports residual language function in chronic post-stroke aphasia. *Hum. Brain Mapp.* **38**, 1636–1658 (2017).
66. Skipper-Kallal, L. M., Lacey, E. H., Xing, S. & Turkeltaub, P. E. Right hemisphere remapping of naming functions depends on lesion size and location in poststroke aphasia. *Neural. Plast.* **2017**, 8740353 (2017).
67. Cardebat, D. et al. Behavioural and neurofunctional changes over time in healthy and aphasic subjects: a PET language activation study. *Stroke* **34**, 2900–2906 (2003).
68. Blank, S. C., Bird, H., Turkheimer, F. & Wise, R. J. Speech production after stroke: the role of the right pars opercularis. *Ann. Neurol.* **54**, 310–320 (2003).
69. Thiel, A. et al. Plasticity of language networks in patients with brain tumours: a positron emission tomography activation study. *Ann. Neurol.* **50**, 620–629 (2001).
70. Crisp, J. & Lambon Ralph, M. A. Unlocking the nature of the phonological-deep dyslexia continuum: the keys to reading aloud are in phonology and semantics. *J. Cogn. Neurosci.* **18**, 348–362 (2006).
71. Hartwigsen, G. et al. Rapid short-term reorganization in the language network. *eLife* **6**, e25964 (2017).
72. Xu, J. S. et al. Task-related concurrent but opposite modulations of overlapping functional networks as revealed by spatial IA. *Neuroimage* **79**, 62–71 (2013).
73. Geranmayeh, F., Leech, R. & Wise, R. J. Network dysfunction predicts speech production after left hemisphere stroke. *Neurology* **86**, 1296–1305 (2016).
74. Southwell, D. G., Hervey-Jumper, S. L., Perry, D. W. & Berger, M. S. Intraoperative mapping during repeat awake craniotomy reveals the functional plasticity of adult cortex. *J. Neurosurg.* **124**, 1460–1469 (2016).
75. Collignon, O. et al. Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. *Brain* **136**, 2769–2783 (2013).
76. Kujala, T. et al. Electrophysiological evidence for cross-modal plasticity in humans with early- and late-onset blindness. *Psychophysiology* **34**, 213–216 (1997).
77. Burton, H. & McLaren, D. G. Visual cortex activation in late-onset, braille naive blind individuals: an fMRI study during semantic and phonological tasks with heard words. *Neurosci. Lett.* **392**, 38–42 (2006).
78. Anderson, C. A., Wiggins, I. M., Kitterick, P. T. & Hartley, D. E. H. Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. *Proc. Natl Acad. Sci. USA* **114**, 10256–10261 (2017).
79. Saur, D. et al. Dynamics of language reorganization after stroke. *Brain* **129**, 1371–1384 (2006).
80. Keidel, J. L., Welbourne, S. R. & Lambon Ralph, M. A. Solving the paradox of the equipotential and modular brain: a neurocomputational model of stroke vs. slow-growing glioma. *Neuropsychologia* **48**, 1716–1724 (2010).
81. Raichle, M. E. & Gusnard, D. A. Appraising the brain's energy budget. *Proc. Natl Acad. Sci. USA* **99**, 10237–10239 (2002).
82. Niven, J. E. Neuronal energy consumption: biophysics, efficiency and evolution. *Curr. Opin. Neurobiol.* **41**, 129–135 (2016).
83. Manring, N. D. & Johnson, R. E. Modelling and designing a variable-displacement open-loop pump. *J. Dyn. Syst. Meas. Control.* **118**, 267–271 (1996).
84. Fearnley, J. M. & Lees, A. J. Ageing and Parkinson disease: substantia nigra regional selectivity. *Brain* **114**, 2283–2301 (1991).
85. Farah, M. J. & McClelland, J. L. A computational model of semantic memory impairment - modality specificity and emergent category specificity. *J. Exp. Psychol. Gen.* **120**, 339–357 (1991).
86. Szaflarski, J. P., Allendorfer, J. B., Banks, C., Vannest, J. & Holland, S. K. Recovered vs. not-recovered from post-stroke aphasia: the contributions from the dominant and non-dominant hemispheres. *Restor. Neurol. Neurosci.* **31**, 347–360 (2013).
87. Heiss, W. D., Kessler, J., Thiel, A., Ghaemi, M. & Karbe, H. Differential capacity of left and right hemispheric areas for compensation of poststroke aphasia. *Ann. Neurol.* **45**, 430–438 (1999).
88. Postman-Caucheteux, W. A. et al. Single-trial fMRI shows contralesional activity linked to overt naming errors in chronic aphasic patients. *J. Cogn. Neurosci.* **22**, 1299–1318 (2010).
89. Szaflarski, J. P. et al. Poststroke aphasia recovery assessed with functional magnetic resonance imaging and a picture identification task. *J. Stroke Cerebrovasc. Dis.* **20**, 336–345 (2011).
90. van Oers, C. A. et al. Contribution of the left and right inferior frontal gyrus in recovery from aphasia. A functional MRI study in stroke patients with preserved hemodynamic responsiveness. *Neuroimage* **49**, 885–893 (2010).
91. Fridriksson, J. Preservation and modulation of specific left hemisphere regions is vital for treated recovery from anomia in stroke. *J. Neurosci.* **30**, 11558–11564 (2010).
92. Rice, G. E., Caswell, H., Moore, P., Lambon Ralph, M. A. & Hoffman, P. Revealing the dynamic modulations that underpin a resilient neural network for semantic cognition: an FMRI investigation in patients with anterior temporal lobe resection. *Cereb. Cortex* **28**, 3004–3016 (2018).
93. Binney, R. J. & Ralph, M. A. L. Using a combination of fMRI and anterior temporal lobe rTMS to measure intrinsic and induced activation changes across the semantic cognition network. *Neuropsychologia* **76**, 170–181 (2015).
94. Jung, J. & Ralph, M. A. L. Mapping the dynamic network interactions underpinning cognition: a cTBS-fMRI study of the flexible adaptive neural system for semantics. *Cereb. Cortex* **26**, 3580–3590 (2016).
95. Welbourne, S. R., Woollams, A. M., Crisp, J. & Ralph, M. A. L. The role of plasticity-related functional reorganization in the explanation of central dyslexias. *Cogn. Neuropsychol.* **28**, 65–108 (2011).
96. Hagoort, P., Wassenaar, M. & Brown, C. Real-time semantic compensation in patients with agrammatic comprehension: electrophysiological evidence for multiple-route plasticity. *Proc. Natl Acad. Sci. USA* **100**, 4340–4345 (2003).
97. Crittenden, B. M., Mitchell, D. J. & Duncan, J. Task encoding across the multiple demand cortex is consistent with a frontoparietal and cingulo-opercular dual networks distinction. *J. Neurosci.* **36**, 6147–6155 (2016).
98. Fedorenko, E., Duncan, J. & Kanwisher, N. Broad domain generality in focal regions of frontal and parietal cortex. *Proc. Natl Acad. Sci. USA* **110**, 16616–16621 (2013).
99. Woolgar, A., Bor, D. & Duncan, J. Global increase in task-related fronto-parietal activity after focal frontal lobe lesion. *J. Cogn. Neurosci.* **25**, 1542–1552 (2013).
100. Murray, L. L. The effects of varying attentional demands on the word retrieval skills of adults with aphasia, right hemisphere brain damage, or no brain damage. *Brain Lang.* **72**, 40–72 (2000).
101. Murray, L. L. Attention and other cognitive deficits in aphasia: presence and relation to language and communication measures. *Am. J. Speech Lang. Pathol.* **21**, s51–s64 (2012).
102. Su, C.-Y., Wuang, Y.-P., Lin, Y.-H. & Su, J.-H. The role of processing speed in post-stroke cognitive dysfunction. *Arch. Clin. Neuropsychol.* **30**, 148–160 (2015).
103. Rajtar-Zembaty, A. et al. Application of the trail making test in the assessment of cognitive flexibility in patients with speech disorders after ischaemic cerebral stroke. *Aktual. Neurol.* **15**, 11–17 (2015).
104. Sharp, D. J., Turkheimer, F. E., Bose, S. K., Scott, S. K. & Wise, R. J. Increased frontoparietal integration after stroke and cognitive recovery. *Ann. Neurol.* **68**, 753–756 (2010).
105. Allendorfer, J. B., Kissela, B. M., Holland, S. K. & Szaflarski, J. P. Different patterns of language activation in post-stroke aphasia are detected by overt and covert versions of the verb generation fMRI task. *Med. Sci. Monit.* **18**, CR135–CR137 (2012).
106. Barlow, T. On a case of double cerebral hemiplegia, with cerebral symmetrical lesions. *Br. Med. J.* **2**, 103–104 (1877).
107. Gowers, W. R. in *A Manual of Diseases of the Nervous System* 2nd edn Vol. 2 (ed. Blakiston, P.) 110–125 (P. Blakiston, Son & Co, 1893).
108. Dunst, B. et al. Neural efficiency as a function of task demands. *Intelligence* **42**, 22–30 (2014).
109. Morcom, A. M. & Henson, R. N. A. Increased prefrontal activity with ageing reflects nonspecific neural responses rather than compensation. *J. Neurosci.* **38**, 7303–7313 (2018).
110. Winhuisen, L. et al. Role of the contralateral inferior frontal gyrus in recovery of language function in poststroke aphasia: a combined repetitive transcranial magnetic stimulation and positron emission tomography study. *Stroke* **36**, 1759–1763 (2005).
111. Leff, A. et al. A physiological change in the homotopic cortex following left posterior temporal lobe infarction. *Ann. Neurol.* **51**, 553–558 (2002).
112. Schofield, T. M. et al. Changes in auditory feedback connections determine the severity of speech processing deficits after stroke. *J. Neurosci.* **32**, 4260–4270 (2012).
113. Lee, Y. S., Zreik, J. T. & Hamilton, R. H. Patterns of neural activity predict picture-naming performance of a patient with chronic aphasia. *Neuropsychologia* **94**, 52–60 (2017).
114. Fischer-Baum, S., Jang, A. & Kajander, D. The cognitive neuroplasticity of reading recovery following chronic stroke: a representational similarity analysis approach. *Neural Plast.* **2017**, 2761913 (2017).
115. Saur, D. et al. Early functional magnetic resonance imaging activations predict language outcome after stroke. *Brain* **133**, 1252–1264 (2010).
116. Tyler, L. K., Wright, P., Randall, B., Marslen-Wilson, W. D. & Stamatakis, E. A. Reorganization of syntactic processing following left-hemisphere brain damage: does right-hemisphere activity preserve function? *Brain* **133**, 3396–3408 (2010).
117. Heiss, W. D. & Thiel, A. A proposed regional hierarchy in recovery of post-stroke aphasia. *Brain Lang.* **98**, 118–123 (2006).
118. Ferbert, A. et al. Interhemispheric inhibition of the human motor cortex. *J. Physiol.* **453**, 525–546 (1992).
119. Corbetta, M., Kincade, M. J., Lewis, C., Snyder, A. Z. & Sapir, A. Neural basis and recovery of spatial attention deficits in spatial neglect. *Nat. Neurosci.* **8**, 1603–1610 (2005).
120. Thiel, A. et al. Effects of noninvasive brain stimulation on language networks and recovery in early poststroke aphasia. *Stroke* **44**, 2240–2246 (2013).
121. Schapiro, A. C., McClelland, J. L., Welbourne, S. R., Rogers, T. T. & Lambon Ralph, M. A. Why bilateral damage is worse than unilateral damage to the brain. *J. Cogn. Neurosci.* **25**, 2107–2123 (2013).
122. Berthier, M. L., Pulvermuller, F., Davila, G., Casares, N. G. & Gutierrez, A. Drug therapy of post-stroke aphasia: a review of current evidence. *Neuropsychol. Rev.* **21**, 302–317 (2011).

123. Castren, E. & Hen, R. Neuronal plasticity and antidepressant actions. *Trends Neurosci.* **36**, 259–267 (2013).
124. Ramanathan, D., Tuszynski, M. H. & Conner, J. M. The basal forebrain cholinergic system is required specifically for behaviourally mediated cortical map plasticity. *J. Neurosci.* **29**, 5992–6000 (2009).
125. Woodhead, Z. V. et al. Auditory training changes temporal lobe connectivity in 'Wernicke's aphasia': a randomized trial. *J. Neurol. Neurosurg. Psychiatry* **88**, 586–594 (2017).
126. Naidech, A. M. et al. Phenytoin exposure is associated with functional and cognitive disability after subarachnoid hemorrhage. *Stroke* **36**, 583–587 (2005).
127. Conroy, P., Sotiropoulou Drosopoulou, C., Humphreys, G. F., Halai, A. D. & Lambon Ralph, M. A. Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia. *Brain* **141**, 1815–1827 (2018).
128. Woodhead, Z. V. J. et al. Randomized trial of iReadMore word reading training and brain stimulation in central alexia. *Brain* **141**, 2127–2141 (2018).
129. Zumbansen, A., Peretz, I. & Hebert, S. Melodic intonation therapy: back to basics for future research. *Front. Neurol.* **5**, 11 (2014).
130. Berthier, M. L. & Pulvermuller, F. Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nat. Rev. Neurol.* **7**, 86–97 (2011).
131. Dignam, J. et al. Intensive versus distributed aphasia therapy: a nonrandomized, parallel-group, dosage-controlled study. *Stroke* **46**, 2206–2211 (2015).
132. Dignam, J. K., Rodriguez, A. D. & Copland, D. A. Evidence for intensive aphasia therapy: consideration of theories from neuroscience and cognitive psychology. *PM R* **8**, 254–267 (2016).
133. Plaut, D. C., McClelland, J. L., Seidenberg, M. S. & Patterson, K. Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol. Rev.* **103**, 56–115 (1996).
134. Bucur, M. & Papagno, C. Are transcranial brain stimulation effects long-lasting in post-stroke aphasia? A comparative systematic review and meta-analysis on naming performance. *Neurosci. Biobehav. Rev.* **102**, 264–289 (2019).
135. Ren, C. L. et al. Effect of low-frequency rTMS on aphasia in stroke patients: a meta-analysis of randomized controlled trials. *PLOS ONE* **9**, e102557 (2014).
136. Wiethoff, S., Hamada, M. & Rothwell, J. C. Variability in response to transcranial direct current stimulation of the motor cortex. *Brain Stimul.* **7**, 468–475 (2014).
137. Lopez-Alonso, V., Cheeran, B., Rio-Rodriguez, D. & Fernandez-del-Olmo, M. Inter-individual variability in response to non-invasive brain stimulation paradigms. *Brain Stimul.* **7**, 372–380 (2014).
138. Sliwinska, M. W. et al. Stimulating multiple-demand cortex enhances vocabulary learning. *J. Neurosci.* **37**, 7606–7618 (2017).
139. Elsner, B., Kugler, J., Pohl, M. & Mehrholz, J. Transcranial direct current stimulation (tDCS) for improving aphasia in patients with aphasia after stroke. *Cochrane Database Syst. Rev.* **5**, CD009760 (2015).
140. Chalela, J. A. et al. Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: a prospective comparison. *Lancet* **369**, 293–298 (2007).
141. O'Brien, P., Sellar, R. J. & Wardlaw, J. M. Fogging on T2-weighted MR after acute ischaemic stroke: how often might this occur and what are the implications? *Neuroradiology* **46**, 635–641 (2004).
142. Wieshmann, U. C. et al. Reduced anisotropy of water diffusion in structural cerebral abnormalities demonstrated with diffusion tensor imaging. *Magn. Reson. Imaging* **17**, 1269–1274 (1999).
143. Gong, G. L. et al. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cereb. Cortex* **19**, 524–536 (2009).
144. Marebwa, B. K. et al. Chronic post-stroke aphasia severity is determined by fragmentation of residual white matter networks. *Sci. Rep.* **7**, 8188 (2017).
145. Ivanova, M. V. et al. Diffusion-tensor imaging of major white matter tracts and their role in language processing in aphasia. *Cortex* **85**, 165–181 (2016).
146. Xing, S., Lacey, E. H., Skipper-Kallal, L. M., Zeng, J. & Turkeltaub, P. E. White matter correlates of auditory comprehension outcomes in chronic post-stroke aphasia. *Front. Neurol.* **8**, 54 (2017).
147. Demeurisse, G. & Capon, A. Brain activation during a linguistic task in conduction aphasia. *Cortex* **27**, 285–294 (1991).
148. Hillis, A. E. et al. Subcortical aphasia and neglect in acute stroke: the role of cortical hypoperfusion. *Brain* **125**, 1094–1104 (2002).
149. Geranmayeh, F., Chau, T. W., Wise, R. J. S., Leech, R. & Hampshire, A. Domain-general subregions of the medial prefrontal cortex contribute to recovery of language after stroke. *Brain* **140**, 1947–1958 (2017).
150. Krainik, A., Hund-Georgiadis, M., Zysset, S. & von Cramon, D. Y. Regional impairment of cerebrovascular reactivity and BOLD signal in adults after stroke. *Stroke* **36**, 1146–1152 (2005).
151. Geranmayeh, F., Wise, R. J., Leech, R. & Murphy, K. Measuring vascular reactivity with breath-holds after stroke: a method to aid interpretation of group-level BOLD signal changes in longitudinal fMRI studies. *Hum. Brain Mapp.* **36**, 1755–1771 (2015).
152. Geranmayeh, F., Brownsett, S. L. & Wise, R. J. Task-induced brain activity in aphasic stroke patients: what is driving recovery? *Brain* **137**, 2632–2648 (2014).
153. Barulli, D. & Stern, Y. Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. *Trends Cogn. Sci.* **17**, 502–509 (2013).
154. Nyberg, L. et al. Age-related and genetic modulation of frontal cortex efficiency. *J. Cogn. Neurosci.* **26**, 746–754 (2014).
155. Herbet, G., Maheu, M., Costi, E., Lafargue, G. & Duffau, H. Mapping neuroplastic potential in brain-damaged patients. *Brain* **139**, 829–844 (2016).
156. Thiel, A. et al. From the left to the right: how the brain compensates progressive loss of language function. *Brain Lang.* **98**, 57–65 (2006).
157. Stern, Y. Cognitive reserve in ageing and Alzheimer disease. *Lancet Neurol.* **11**, 1006–1012 (2012).
158. Puente, A. N., Lindbergh, C. A. & Miller, L. S. The relationship between cognitive reserve and functional ability is mediated by executive functioning in older adults. *Clin. Neuropsychol.* **29**, 67–81 (2015).
159. Uiterwijk, R. et al. Total cerebral small vessel disease MRI score is associated with cognitive decline in executive function in patients with hypertension. *Front. Ageing Neurosci.* **8**, 301 (2016).
160. Molad, J. et al. Only white matter hyperintensities predicts post-stroke cognitive performances among cerebral small vessel disease markers: results from the TABASCO study. *J. Alzheimers Dis.* **56**, 1293–1299 (2017).
161. Woollams, A. M., Madrid, G. & Lambon Ralph, M. A. Using neurostimulation to understand the impact of pre-morbid individual differences on post-lesion outcomes. *Proc. Natl Acad. Sci. USA* **114**, 12279–12284 (2017).

Acknowledgements
J.D.S. is a Wellcome Clinical PhD Fellow funded by grant 203914/Z/16/Z, awarded to the universities of Manchester, Leeds, Newcastle and Sheffield, UK. The authors' research is supported by a European Research Council Advanced Grant to M.A.L.R. (GAP: 670428) and a Rosetrees Trust grant to A.H. and M.A.L.R. (A1699).

Author contributions
All authors were involved in the conceptualization and writing of the manuscript.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Neurology thanks A. Hillis, S. Kiran and A. Leff for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Chapter 3

Language networks in aphasia and health: A 1000 participant activation likelihood estimation meta-analysis

Published paper: Stefaniak, J. D., Alyahya, R. S. W., & Lambon Ralph, M. A. (2021). Language networks in aphasia and health: A 1000 participant activation likelihood estimation meta-analysis. *NeuroImage*, 233, 117960.

NeuroImage 233 (2021) 117960

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Language networks in aphasia and health: A 1000 participant activation likelihood estimation meta-analysis

James D. Stefaniak[a,b,c,1,*], Reem S.W. Alyahya[a,d,e,1], Matthew A. Lambon Ralph[a]

[a] MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK
[b] Department of Psychiatry, University of Cambridge, Cambridge, UK
[c] Division of Neuroscience and Experimental Psychology, University of Manchester, Manchester, UK
[d] King Fahad Medical City, Riyadh, Saudi Arabia
[e] Alfaisal University, Riyadh, Saudi Arabia

ARTICLE INFO

Keywords:
Meta-analysis
fMRI
PET
Plasticity
Language
Stroke

ABSTRACT

Aphasia recovery post-stroke is classically and most commonly hypothesised to rely on regions that were not involved in language premorbidly, through 'neurocomputational invasion' or engagement of 'quiescent homologues'. Contemporary accounts have suggested, instead, that recovery might be supported by under-utilised areas of the premorbid language network, which are downregulated in health to save neural resources ('variable neurodisplacement'). Despite the importance of understanding the neural bases of language recovery clinically and theoretically, there is no consensus as to which specific regions are more likely to be activated in post-stroke aphasia (PSA) than healthy individuals. Accordingly, we performed an Activation Likelihood Estimation (ALE) meta-analysis of language functional neuroimaging studies in PSA. We obtained coordinate-based functional neuroimaging data for 481 individuals with aphasia following left-hemisphere stroke and 530 linked controls from 33 studies that met predefined inclusion criteria. ALE identified regions of consistent, above-chance spatial convergence of activation, as well as regions of significantly different activation likelihood, between participant groups and language tasks. Overall, these findings dispute the prevailing theory that aphasia recovery involves recruitment of novel right hemisphere territory into the language network post-stroke. Instead, multiple regions throughout both hemispheres were consistently activated during language tasks in both PSA and controls. Regions of the right anterior insula, frontal operculum and inferior frontal gyrus (IFG) pars opercularis were more likely to be activated across all language tasks in PSA than controls. Similar regions were more likely to be activated during higher than lower demand comprehension or production tasks, consistent with them representing enhanced utilisation of spare capacity within right hemisphere executive-control related regions. This provides novel evidence that 'variable neurodisplacement' underlies language network changes that occur post-stroke. Conversely, multiple undamaged regions were less likely to be activated across all language tasks in PSA than controls, including domain-general regions of medial superior frontal and paracingulate cortex, right IFG pars triangularis and temporal pole. These changes might represent functional diaschisis, and demonstrate that there is not global, undifferentiated upregulation of all domain-general neural resources during language in PSA. Such knowledge is essential if we are to design neurobiologically-informed therapeutic interventions to facilitate language recovery.

Abbreviations

ALE	Activation Likelihood Estimation
FWHM	Full Width at Half Maximum
FWE	Family Wise Error
fMRI	functional MRI
IFG	Inferior Frontal Gyrus
MCA	Middle Cerebral Artery
MA	Modelled Activation
MD	Multiple Demand
MFG	Middle Frontal Gyrus
MNI	Montreal Neurological Institute
MTG	Middle Temporal Gyrus
PSA	Post-Stroke Aphasia
ROI	Region Of Interest
SFG	Superior Frontal Gyrus
SMC	Supplementary Motor Cortex
STG	Superior Temporal Gyrus

* Corresponding author at: MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK
E-mail address: james.stefaniak@cantab.net (J.D. Stefaniak).
[1] Joint first authors.

https://doi.org/10.1016/j.neuroimage.2021.117960
Received 28 October 2020; Received in revised form 28 February 2021; Accepted 8 March 2021
Available online 17 March 2021

SVC Small Volume Correction.

1. Introduction

Post-stroke aphasia (PSA) is prevalent and debilitating (Engelter et al., 2006) and recovery of function tends to be variable and often incomplete (Yagata et al., 2017). Compensatory changes in patterns of neural activity, reflecting increased utilisation of surviving neural regions, are hypothesised to contribute to aphasia recovery (Murphy and Corbett, 2009; Stefaniak et al., 2020; Turkeltaub et al., 2011). While previous studies have explored which set of regions are consistently activated in PSA (Turkeltaub et al., 2011), multiple key questions remain unanswered. These include: (a) which regions, if any, are more or less likely to be activated in PSA than healthy individuals across all language tasks and do these regions differ between language tasks of different nature (comprehension vs. production); (b) are regions upregulated in PSA also modulated by task difficulty (higher vs. lower demand); and (c) do the differentially activated regions vary between different stages of recovery. Such knowledge will be essential to understand the mechanisms underlying language network plasticity and thus design neurobiologically-informed therapeutic interventions to aid language recovery. Accordingly, this study tackled these targeted questions through the largest Activation Likelihood Estimation (ALE) meta-analysis, to date, of functional neuroimaging studies in PSA (n=481) and healthy controls (n=530). We define the language network as regions consistently activated during language, which might include both language-specific regions, reportedly activated during language but not non-language tasks (Fedorenko et al., 2011; Pritchett et al., 2018), as well as domain-general regions activated during both language and non-language tasks (Fedorenko et al., 2013; Geranmayeh et al., 2017). There were several specific questions we sought to address. We consider these briefly, below, with respect to three major themes.

First, even though recovery of language after stroke has perplexed researchers since the seminal studies of aphasia in the nineteenth century (Finger et al., 2003), there have been very few formal, implemented models (Chang and Lambon Ralph, 2020) and hypotheses have rarely been tested in relation to large patient datasets (Stefaniak et al., 2020). Certain mechanisms underlying partial language recovery in PSA propose that neural networks unused during language in health can adapt after stroke to perform a similar function to the one normally supported by the now damaged neural network(s) (Stefaniak et al., 2020), for instance through immediate engagement of *quiescent homologues* (Finger et al., 2003) or through *neurocomputational invasion* of non-language regions via experience-dependent plasticity (Keidel et al., 2010; Southwell et al., 2016). Alternatively, *variable neurodisplacement* (Binney and Lambon Ralph, 2015; Stefaniak et al., 2020) proposes that 'well engineered' language and cognitive networks dynamically balance performance demand against energy expenditure, downregulating spare capacity under standard performance demands in health but running the remaining system 'harder' after partial damage (as the intact system can do when under increased performance demands (Jung and Lambon Ralph, 2016; Rice et al., 2018; Robson et al., 2014; Sharp et al., 2010)). These mechanisms are not mutually exclusive and might include both language-specific (Fedorenko et al., 2011; Pritchett et al., 2018) and non-language networks, including domain-general executive networks (Fedorenko et al., 2013), in both hemispheres (Stefaniak et al., 2020). Key predictions of *variable neurodisplacement* are that compensatory language network changes in PSA are due to upregulation of spare capacity within the pre-existing language network, and that these same upregulated neural regions show increased activation for hard over easier tasks in both PSA and healthy individuals.

Second, there is a tendency to treat 'language' and its recovery as a single, homogenous cognitive function. Instead, language refers to a diverse range of expressive and receptive activities. Different language activities are supported by interactions between various more general neurocognitive computations (Gordon et al., 2002; Mementi et al., 2011; Patterson and Lambon Ralph, 1999) which can be damaged independently of each other to generate the graded, multidimensional nature of post-stroke aphasia (Alyahya et al., 2020; Butler et al., 2014; Kummerer et al., 2013; Mirman et al., 2015). Consequently, theories of recovery need to consider not only how each primary neurocognitive system might recover, but also how changes in their interactivity can support improved performance across different language activities. Changes in the division of labour across systems can occur not only between language networks (Ueno et al., 2011) but also between language and multi-demand executive systems (Geranmayeh et al., 2017; Hartwigsen, 2018).

An important second aspect of this issue is that different subcomponents of language, such as those subserving comprehension versus production, might have differently distributed networks, including degrees of lateralisation, premorbidly (Lidzba et al., 2011). For instance, the language network is often described as unilateral (Mazoyer et al., 2014) but several lines of evidence suggest it is at least partially bilateral but asymmetric (Fedorenko et al., 2011; Lambon Ralph et al., 2001). This has significant implications as many studies have highlighted a role for the right hemisphere in recovery (Crinion and Price, 2005; Skipper-Kallal et al., 2017a, b). Depending on the degree of premorbid asymmetry, right hemisphere activation might reflect engagement of pre-existing right hemispheric regions of the language network via *variable neurodisplacement* versus novel recruitment of non-language regions via *neurocomputational invasion* (Chang and Lambon Ralph, 2020; Warburton et al., 1999). It is important, therefore, to compare activation patterns in post-stroke aphasia with the natural distribution of the same language subcomponent(s) in healthy individuals.

Third, language recovery is dynamic and occurs most rapidly during the first few months post-stroke (Pedersen et al., 1995; Yagata et al., 2017), with spontaneous language changes being slower and smaller by the 'chronic' stage after approximately 6-12 months (Hope et al., 2017). Thus, in order to identify language network changes that are associated with recovery, it is important to compare language networks at subacute vs. chronic stages of recovery.

Given these many outstanding questions, this study sought to identify regions of consistent, above-chance spatial convergence of activation, as well as regions of significantly different activation likelihood, between participant groups and language tasks. The omnibus ALE meta-analysis considered which specific regions are more likely to be activated in PSA than healthy individuals across all language tasks. Subsequent subgroup analyses investigated differences based on: comprehension versus production tasks; for each task type, higher versus lower demand tasks; and time post stroke (i.e., sub-acute vs. chronic PSA). Unfortunately, there were too few studies of sub-acute patients in the literature to contrast them against chronic PSA in this meta-analysis. If language recovery reflects *neurocomputational invasion* or engagement of *quiescent homologues* then the post-stroke language network should expand to include novel regions that are not consistently activated in healthy individuals, even under increased task difficulty. Conversely, *variable neurodisplacement* predicts that the networks observed in PSA should also be observed in healthy controls, particularly when the healthy system is placed under greater performance demands.

2. Materials and methods

2.1. Study search and selection

We searched the databases Medline, Embase and PsycINFO up to April 2020. Terms relating to aphasia (aphasia OR dysphasia OR language OR fluency OR phonology OR semantics OR naming OR repetition OR comprehension OR speaking), stroke (stroke OR ischaemia OR ischemia OR infarct) and neuroimaging (fMRI OR PET OR neuroimaging OR imaging OR functional) were used. We identified eligible articles reporting observational studies that had: a) more than one person with

language impairment at any time following a single left hemispheric stroke; b) more than one healthy control; and c) performed fMRI or ^{15}O-PET during language task-based functional neuroimaging. We extracted coordinate data for inclusion in this ALE meta-analysis that: related to activation (not deactivation) during a language task-based functional neuroimaging experiment; was provided in standard space; was derived from whole-brain mass-univariate analyses without region of interests (ROIs), small volume corrections (SVC), or conjunctions (Müller et al., 2018); was reported separately for PSA and control groups; and was calculated using the same significance thresholds in the PSA and control groups. We excluded coordinate data from survivors of right hemisphere strokes or with multiple previous strokes. Full details are reported in the Supplementary Information. If coordinates meeting these criteria for both the PSA and control groups were not provided in the publication, the authors were contacted to request unpublished coordinates.

2.2. ALE meta-analysis

Peak coordinates pertaining to language activation were extracted from each included article and double checked by the same author (JDS). Coordinates in Talairach space were converted to Montreal Neurological Institute (MNI) space using the Lancaster transformation (Lancaster et al., 2007). GingerALE 3.0.2 was used to perform ALE (http://brainmap.org/ale/), which is a random-effects coordinate-based meta-analytic technique that identifies neural regions at which activation peaks converge above-chance across participant groups within a single dataset (Eickhoff et al., 2012; Eickhoff et al., 2011; Eickhoff et al., 2009; Turkeltaub et al., 2012). Individual studies might have reported activation coordinates for multiple subgroups of participants and thus contributed more than one participant group to the meta-analysis. Briefly, we grouped together activation peaks from all imaging tasks performed by the same participant group. Each peak was modelled as a 3D Gaussian distribution of activation probability with a Full Width at Half Maximum (FWHM) based on empirical estimates of spatial uncertainty derived from the number of participants in the group, with larger sample sizes modelled by narrower, taller Gaussians providing a more reliable approximation of the true activation location (Eickhoff et al., 2009). Each voxel within a default grey matter mask was assigned the activation probability from the peak within the shortest Euclidean distance, producing a Modelled Activation (MA)-map for each participant group (Turkeltaub et al., 2012). The voxel-wise union of all MA-maps from all participant groups included in a single dataset produced an ALE-map, in which ALE values represent the likelihood that at least one participant group activated a given voxel (Turkeltaub et al., 2012). For single dataset analyses, we tested the null hypothesis of random spatial association between participant groups ('spatial independence'), namely that any spatial convergence of activation between different participant groups in a dataset is only occurring by chance (Eickhoff et al., 2012). In order to compute, analytically, the null distribution of ALE values under the assumption of spatial independence between participant groups, each participant group's MA-map was first converted into a histogram representing the probability of observing each MA value in that map (Eickhoff et al., 2012). Histograms representing MA-maps of individual participant groups were iteratively combined (Eickhoff et al., 2012) to produce a final histogram representing the probability of observing any given ALE value under the null hypothesis of spatial independence between participant groups. The null distribution and ALE-map were combined to produce a p-value map for each dataset. The p-value map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error (FWE) corrected threshold of $p<0.05$ based on 1000 random permutations (Eickhoff et al., 2016). Briefly, the null distribution of cluster sizes given a cluster-forming threshold of $p<0.001$ was obtained through random simulation in which, for every participant group, a matched simulated participant group was created containing the same number of participants and foci but with foci randomly located throughout the grey matter mask (Eickhoff et al., 2012). The above ALE meta-analytical algorithm was performed on each simulated dataset and each simulated ALE-map thresholded at the cluster-forming threshold of $p<0.001$. The size of each contiguous cluster of suprathreshold voxels was recorded for each of 1000 such randomly simulated ALE-maps to produce a distribution of cluster sizes that would be expected under the null hypothesis of spatial independence between participant groups (Eickhoff et al., 2012). Suprathreshold clusters in the real dataset's ALE-map that were larger than 95% of the null distribution clusters were significant at FWE $p<0.05$ and taken to represent regions in which spatial convergence of activation between different participant groups was significantly above chance, which we define herein as regions of consistent activation.

Coordinates from tasks at different timepoints on the same participant group were not pooled; only tasks performed at the longest timepoint post-stroke for each group were included. If coordinates were available for separate groups within the same study (e.g., for stroke survivors with aphasia as individuals or sub-groups), each individual/subgroup was counted as being from a separate participant group in the meta-analysis. Single participants were included as 'participant groups' of size $n=1$; as explained above, the FWHM of the Gaussian probability distribution of each peak was weighted to take account of the increasing spatial uncertainty associated with decreasing group size (Eickhoff et al., 2009).

Conjunction images identifying regions in which two datasets both showed consistent activation were computed as the intersection of the thresholded ALE-maps (Eickhoff et al., 2011). Contrast analyses were performed to identify regions where activation likelihood differed significantly between two datasets. ALE-maps from the two datasets being contrasted were subtracted from each other and thresholded at $p<0.05$ (uncorrected) using 10,000 *P*-value permutations with a minimum cluster threshold of 200mm^3. Each permutation involved pooling all participant groups contributing to either dataset alone and randomly dividing them into two datasets of the same size (i.e. number of participant groups) as the two original datasets being contrasted (Eickhoff et al., 2011). ALE-maps for these two randomly assembled datasets were calculated and the difference between these 'random' ALE-maps computed. Repeating this 10,000 times produced a null-distribution for the difference in ALE values between the two datasets expected under the null hypothesis of label exchangeability at each voxel in the brain (Eickhoff et al., 2011). The observed difference in ALE values at each voxel was compared to its null distribution, yielding a p-value map that was thresholded at $p<0.05$ with a minimum cluster threshold of 200 mm^3 and inclusively masked to voxels that were significant during single dataset meta-analysis of either included dataset (Eickhoff et al., 2011). This method of permutation testing accounted for differences in the number of participant groups between each dataset being contrasted.

The Harvard-Oxford atlas (Desikan et al., 2006) defined anatomical labels and the Talairach Daemon atlas (Lancaster et al., 2000) determined the Brodmann Area label associated with each peak coordinate.

We performed a set of pre-planned ALE meta-analyses that are set out below. For the omnibus ALE meta-analysis comparing all language tasks between PSA and controls groups, we required single datasets to have at least 17 participant groups, as recommended by empirical simulations suggesting this number was needed to ensure adequate power (Eickhoff et al., 2016). Given the scarcity of functional neuroimaging studies in PSA, we required 10 participant groups for single datasets to be included in ALE meta-analyses for more specific contrasts between subgroups of participants or tasks, as per previous recommendations (Eickhoff and Bzdok, 2013). Single datasets never contained data from the same participants as separate participant groups. If the same participant group performed multiple imaging tasks which were divided into different datasets during contrast analyses (e.g. both higher and lower demand comprehension tasks), the coordinates for both imaging tasks were included in their respective datasets. Since contrast subgroup analyses were designed to look for regions of significantly different acti-

vation likelihood between groups, the inclusion of coordinates from the same group in both subgroup datasets being contrasted would, if anything, reduce the likelihood of finding differences and thus should not increase the false positive rate.

2.3. Differences between PSA and control groups

2.3.1. All language tasks in PSA vs. controls (omnibus analysis)

This analysis combined all data available. Thus, it consisted of single dataset, conjunction and contrast ALE meta-analyses comparing all language tasks in all PSA against all language tasks in all controls. The included coordinates did not contain duplicated data from the same participants.

2.3.2. Comprehension and production tasks in PSA vs. controls

PSA participants might activate different neural regions relative to controls for a subset of language tasks. Such differences may have been obscured by grouping all language tasks together in the omnibus ALE meta-analysis. Participant groups were therefore divided according to whether their functional neuroimaging tasks involved 'production' (including either overt or covert production of sublexical, lexical or sentence level speech components) or solely 'comprehension' without production (e.g. sentence listening, semantic judgement, picture-word matching). Single dataset, conjunction and contrast ALE meta-analyses were conducted to compare comprehension tasks in PSA against controls, and production tasks in PSA against controls.

2.3.3. Comprehension>production and production>comprehension tasks in PSA vs. controls

Changes in the division of labour between networks subserving distinct underlying language functions might support improved language performance post-stroke. Conjunction and contrast ALE meta-analyses were performed to compare comprehension vs. production tasks, separately within PSA and control groups. Significant clusters for 'comprehension>production' and 'production>comprehension' were then qualitatively compared between PSA and control groups.

2.4. Higher versus lower processing demand tasks

Variable neurodisplacement proposes that neural spare capacity is downregulated to save energy under standard performance demands in health but is upregulated when performance demands increase post-stroke. If this occurs, we would expect the neural regions upregulated in PSA to be more likely to be activated during more difficult compared to less difficult tasks in both PSA and controls. Therefore, comprehension and production tasks were each subdivided according to task difficulty. Higher demand comprehension tasks were defined as tasks requiring a linguistic decision to be made; e.g., whether a stimulus is a word or pseudoword, concrete or abstract, or related to some other semantic or syntactic property. Lower demand comprehension tasks either did not require a linguistic decision or required a very simple identity match; e.g., passive listening or simple word-picture matching. Higher demand production tasks required production of >1 word, such as propositional speech or category fluency tasks. Lower demand production tasks required production of single words, such as picture naming or single item repetition. Single dataset, conjunction and contrast ALE meta-analyses were conducted to compare higher versus lower demand comprehension tasks, and higher versus lower demand production tasks. These contrasts were initially performed separately within PSA and control groups. However, there were too few participant groups to contrast higher versus lower processing demand comprehension or production tasks in controls, so a third set of analyses combined PSA and control participant groups together. Significant clusters representing demand-responsive regions were compared to regions of significantly different activation likelihood between PSA and control groups identified by the meta-analyses in Section 2.3.

Clusters identified in the above analyses were also compared for spatial overlap with the Multiple Demand (MD) network (Duncan, 2010), a set of domain-general neural regions activated during a diverse range of executively demanding language and non-language cognitive tasks (Fedorenko et al., 2013), and with the semantic control network known to be involved during executively demanding semantic cognition in healthy individuals (Jackson, 2021).

2.5. Time post-stroke

Language recovery occurs most rapidly during the first six months post-stroke (Pedersen et al., 1995; Yagata et al., 2017). PSA groups were therefore categorised according to whether their mean time post-stroke was before or after 6 months. Unfortunately, there were too few studies of sub-acute patients to contrast them formally with chronic PSA.

2.6. Statistical analysis

We compared mean ages of the PSA and control groups using Mann-Whitney U tests implemented in SPSS version 25 with statistical significance defined as $p<0.05$ with Bonferroni correction.

2.7. Data availability

Group level coordinate data supporting the findings of this study are available on figshare (doi: 10.6084/m9.figshare.12582935).

3. Results

3.1. Descriptive statistics

10,169 unique references were obtained from the systematic search. 79 papers were eligible for inclusion; useable foci were obtained from 33/79 included papers. A flowchart of the search and selection process is shown in Fig. 1. Details of the included/excluded papers, reasons for excluding eligible papers, and information on the PSA groups included in the ALE meta-analysis are provided in Supplementary Tables S1-3. Across all language tasks, 1521 foci were obtained from 481 PSA in 64 groups, and 809 foci were obtained from 530 healthy controls in 37 groups (Supplementary Tables S3, 4). Foci relating to 172 of the 481 PSA had not been published but were provided after personal communication with the corresponding authors (Barbieri et al., 2019; Geranmayeh et al., 2016; Hallam et al., 2018; Meier et al., 2019; Radman et al., 2016; Schofield et al., 2012; Tao and Rapp, 2019; Wilson et al., 2018).

The 64 PSA groups did not have significantly different mean ages compared to the 37 control groups (median 57.4 [IQR 9.0] years in PSA groups vs. 57.0 [IQR 8.2] years in control groups; Mann-Whitney U-test, U=878, two-sided $p=0.18$). Every pair of datasets contrasted in this paper had mean ages that were not statistically significantly different (Supplementary Table S33). Fig. 2 contains histograms of the mean ages of the groups.

3.2. Differences between PSA and control groups

Our first aim was to investigate which, if any, regions are more or less likely to be activated in PSA than healthy individuals across all language tasks and do these regions differ between language tasks of different nature (comprehension vs. production).

3.2.1. All language tasks in PSA vs. controls (omnibus analysis)

Single datasets from the omnibus meta-analysis comparing all language tasks in all PSA against control groups are reported in the Supplementary Information and illustrated in Fig. 3. A conjunction demonstrated that both PSA and control groups consistently activated overlapping regions in: left frontal lobe (frontal operculum cortex, IFG

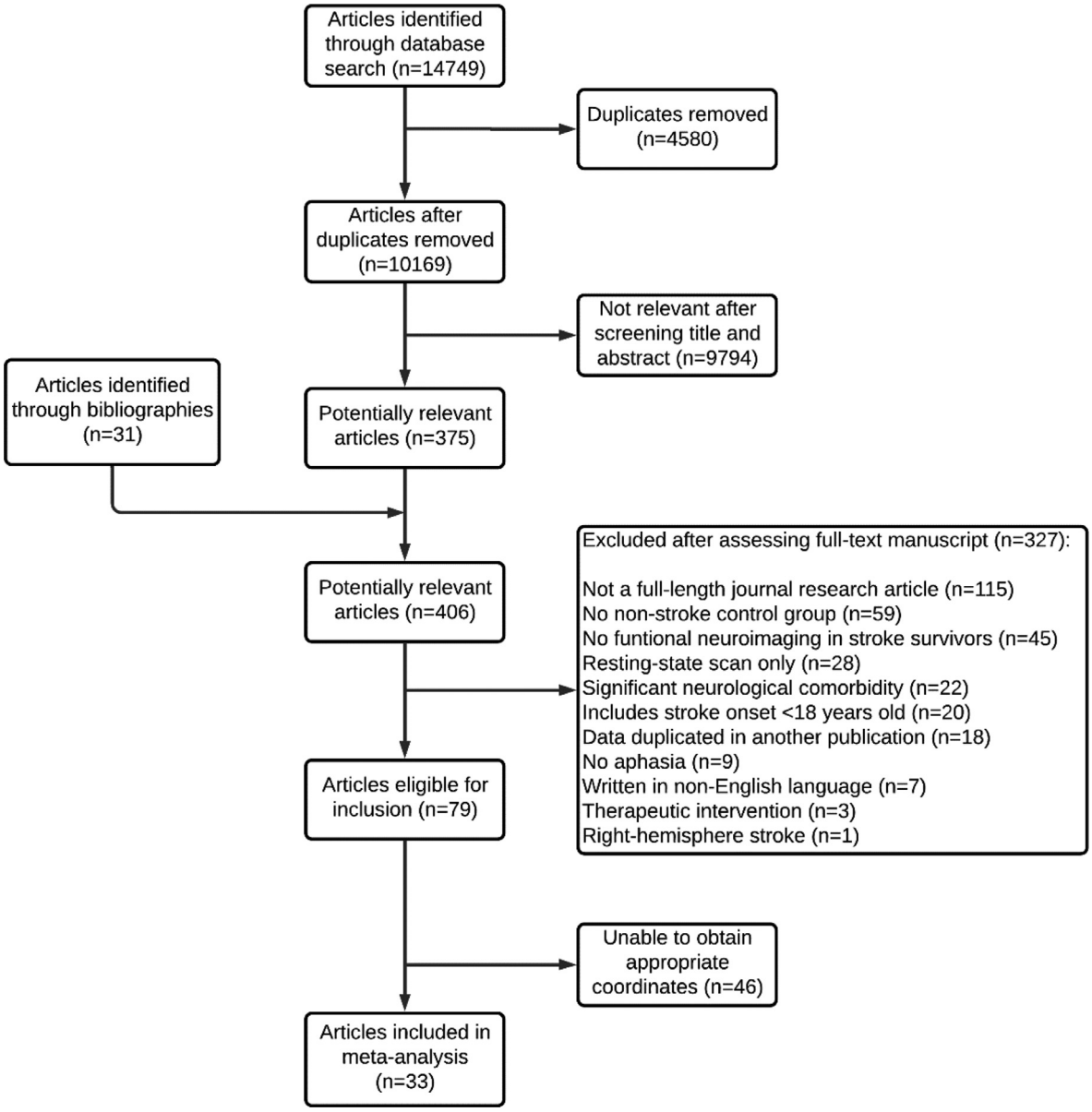


Fig. 1. Flowchart of the selection process for included papers. Flowchart showing the selection process at each stage of the systematic search up to April 2020. Ultimately, activation foci from 33 papers were included in the ALE meta-analysis.

pars opercularis/triangularis, frontal orbital cortex, MFG); left posterior MTG; midline cortex (SFG, SMC, paracingulate gyrus); right frontal lobe (frontal operculum, frontal orbital cortex); right posterior STG; and right posterior supramarginal gyrus (Supplementary Table S7). This highlights that multiple regions throughout both hemispheres were consistently activated in PSA but were also involved in language pre-morbidly rather than being recruited 'de novo' post-stroke. Conjunction clusters in the left frontal lobe (frontal operculum cortex, IFG pars opercularis/triangularis, MFG), midline cortex (SFG, SMC, paracingulate cortex) and right frontal lobe (frontal operculum, frontal orbital cortex) at least partially overlap with the MD network (Fedorenko et al., 2013), suggesting that the language network includes domain-general regions in both controls and PSA.

Contrast analyses revealed that multiple regions were less likely to be activated during language in the PSA group than controls, including midline SFG, SMC, and paracingulate gyrus as well as right IFG pars triangularis and right temporal pole (Supplementary Table S8). The midline SFG and paracingulate gyrus cluster overlaps with the MD network (Fedorenko et al., 2013), suggesting it is domain-general in controls (Fig. 6A), whereas the right IFG pars triangularis and right temporal pole clusters do not. Since all strokes were restricted to the left hemisphere, this result demonstrates that a set of undamaged language and domain-general regions are less likely to be activated in PSA than controls.

The PSA group were more likely to activate the right anterior insula, frontal operculum and IFG pars opercularis during language than controls (Supplementary Table S8). This cluster overlaps with the Multiple Demand (MD) network (Fedorenko et al., 2013) in the right frontal operculum and anterior insula, suggesting it is domain-general in controls (Fig. 6A). Parts of the right anterior insula, frontal operculum and IFG pars opercularis were consistently activated across all language tasks in controls (Supplementary Table S6).

3.2.2. *Comprehension tasks in PSA vs. controls*

Single datasets from the subgroup meta-analysis comparing comprehension tasks in all PSA against control groups are reported in the

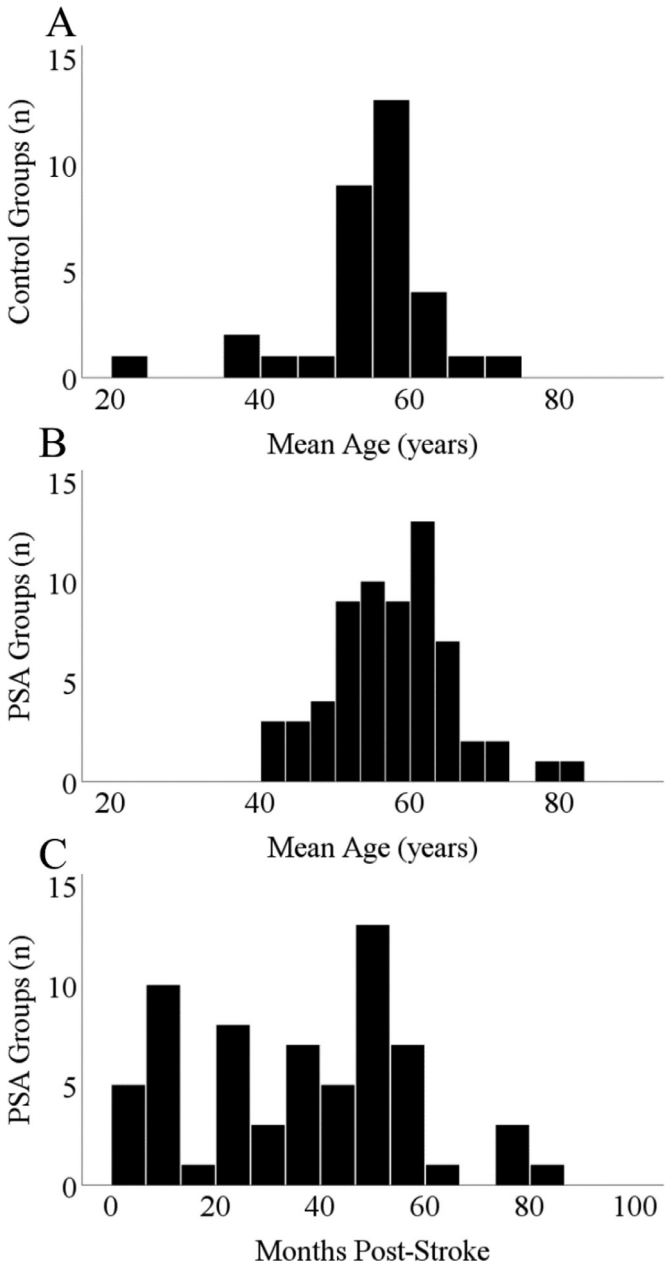


Fig. 2. Histogram showing the distribution of participant groups with age and time post-stroke. A: Histogram showing the number of control groups for each 'mean age'. B: Histogram showing the number of PSA groups for each 'mean age'. C: Histogram showing the number of PSA groups for each post-stroke time period.

Supplementary Information and illustrated in Fig. 4. A conjunction demonstrated that both PSA and controls consistently activated overlapping regions during comprehension in left frontal lobe (Fig. 4A, IFG pars opercularis/triangularis, frontal orbital cortex) and left posterior MTG (Supplementary Table S11).

Contrast analyses revealed that multiple regions were less likely to be activated during comprehension in the PSA group than controls, including midline cortical regions (SFG, paracingulate gyrus) that are unlikely to be damaged following a middle cerebral artery (MCA) stroke (Fig. 4B, Supplementary Table S12). This midline SFG/paracingulate gyrus cluster does not overlap with the MD network (Fig. 6B) (Fedorenko et al., 2013). PSA were more likely to activate the right anterior insula and frontal operculum during comprehension than controls (Fig. 4B, Supplementary Table S12); this cluster overlaps with the MD network (Fedorenko et al., 2013), suggesting it is domain-general in controls (Fig. 6B).

3.2.3. Production tasks in PSA vs. controls

Single datasets from the subgroup meta-analysis comparing production tasks in all PSA against control groups are reported in the Supplementary Information and illustrated in Fig. 4. A conjunction demonstrated that both PSA and controls consistently activated overlapping regions during production in: left IFG pars triangularis; midline cortex (SFG, SMC, paracingulate gyrus); and right posterior STG (Fig. 4C, Supplementary Table S15). This highlights that multiple regions throughout both hemispheres are consistently activated during language production in PSA that were involved in language pre-morbidly rather than being recruited 'de novo' post-stroke. Conjunction clusters in the midline SFG, SMC and paracingulate gyrus overlap with the MD network (Fedorenko et al., 2013), suggesting that the language production network includes domain-general regions in both controls and PSA.

Contrast analyses revealed that PSA were less likely than controls to activate the following midline and right hemisphere regions during production: midline cortex (SFG, SMC, paracingulate gyrus); right frontal lobe (frontal orbital cortex, precentral gyrus); right insula; and right temporal lobe (Heschl's gyrus, posterior STG, temporal pole) (Fig. 4D, Supplementary Table S16). Again, these regions fall outside of the left MCA territory and thus were unlikely to have been lesioned by the stroke. The midline SFG/SMC/paracingulate gyrus, right frontal orbital cortex and anterior insula clusters overlap with the MD network (Fedorenko et al., 2013), suggesting they are domain-general in controls (Fig. 6C). No regions were more likely to be activated during production in the PSA group than controls (Supplementary Table S16).

3.2.4. Comprehension>Production tasks in PSA vs. controls

Changes in the division of labour between networks subserving distinct underlying language functions might support improved language performance post-stroke (Stefaniak et al., 2020).

A conjunction demonstrated that controls consistently activated overlapping regions during both comprehension and production tasks in: left IFG pars opercularis/triangularis; and left temporal lobe (posterior MTG, temporooccipital MTG) (Supplementary Table S19). Compared to controls, PSA had additional clusters of conjunction during both comprehension and production tasks in the midline cortex (SFG, SMC) and right frontal lobe (frontal operculum cortex, frontal orbital cortex) (Supplementary Table S17).

Contrast analyses revealed that the left frontal lobe (frontal orbital cortex, frontal pole), left temporal lobe (temporal pole, temporooccipital inferior temporal gyrus) and midline SFG/paracingulate gyrus were significantly more likely to be activated during comprehension than production in controls (Fig. 4E, Supplementary Table S20). PSA had additional clusters of increased activation likelihood during comprehension than production in the right anterior insula and right MFG that were not observed in controls (Fig. 4F, see Supplementary Information and Supplementary Table S18 for full details). These two PSA-specific clusters overlap with the semantic control network (Jackson, 2021) (Fig. 7A) and with the MD network (Fedorenko et al., 2013) (Fig. 6D). The right anterior insula cluster overlaps with the region of greater activation likelihood in PSA than controls during comprehension tasks. Taken together, these results suggest that comprehension tasks in PSA make greater use of specific right frontal domain-general regions than both production tasks in PSA (right anterior insula, MFG), and comprehension tasks in controls (right anterior insula).

3.2.5. Production>Comprehension tasks in PSA vs. controls

Contrast analyses revealed that the left frontal lobe (IFG pars opercularis/triangularis, frontal orbital cortex, precentral gyrus), left insula, left temporal lobe (planum temporale, temporooccipital MTG), left

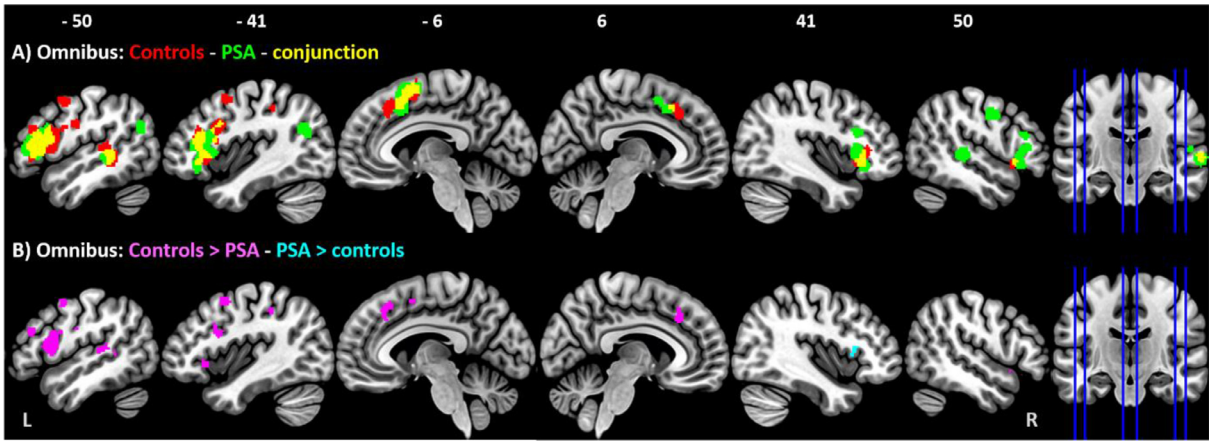


Fig. 3. Omnibus ALE meta-analysis for all language tasks in PSA and healthy controls. A: ALE maps of all tasks in PSA (green clusters), in controls (red clusters) and conjunction map of all tasks in both PSA and controls (yellow clusters). ALE single dataset analyses thresholded at $p<0.001$ uncorrected voxel-wise, FWE $p<0.05$ cluster wise, 1000 permutations. B: ALE maps of 'all language tasks: controls > PSA' (violet clusters) and 'all language tasks: PSA > controls' (cyan clusters). ALE contrast analyses thresholded at $p<0.05$, 10000 permutations, minimum cluster extent 200ml.

posterior supramarginal gyrus, midline cortex (SFG, SMC, paracingulate gyrus) and right temporal lobe (temporal pole, posterior STG) had greater activation likelihood during production than comprehension in controls (Fig. 4E, Supplementary Table S20). PSA had an additional cluster of increased activation likelihood during production than comprehension in the right precentral gyrus that was not observed in controls (Fig. 4F, see Supplementary Information and Supplementary Table S18 for full details). This PSA-specific right precentral gyrus cluster did not overlap with the MD network (Fedorenko et al., 2013). These results suggest that production tasks in PSA make greater use of a specific right precentral gyrus region than comprehension tasks in PSA, but this differential activation was not present in controls.

3.2.6. Summary

Multiple regions throughout both hemispheres, including domain-general regions, are consistently activated during language in both PSA and controls. PSA are more likely to activate the following regions than controls: right anterior insula, right frontal operculum (all language tasks, comprehension tasks) and right IFG pars opercularis (all language tasks). PSA are less likely to activate the following regions than controls: midline SFG/SMC/paracingulate gyrus (all language tasks, comprehension tasks, production tasks); right IFG pars triangularis (all language tasks); right frontal orbital cortex, precentral gyrus, anterior insula, Heschl's gyrus, posterior STG (production tasks); and right temporal pole (all language tasks, production tasks). The networks subserving comprehension vs. production tasks diverge in PSA relative to controls. Comprehension tasks in PSA make greater use of specific right frontal regions than both production tasks in PSA (right anterior insula, MFG), and comprehension tasks in controls (right anterior insula). Conversely, production tasks in PSA make greater use of a right precentral gyrus region than comprehension tasks in PSA, but this differential activation was not present in controls.

3.3. Regions modulated by task difficulty

Our second aim was to investigate whether regions upregulated in PSA are also modulated by task difficulty, in keeping with *variable neurodisplacement*.

3.3.1. Higher versus lower demand comprehension tasks

Single datasets from the meta-analysis comparing higher vs. lower demand comprehension tasks in PSA are reported in the Supplementary Information.

Contrast analyses revealed that clusters in the left frontal lobe (IFG pars opercularis/triangularis, MFG), right frontal lobe (frontal operculum cortex, IFG pars opercularis/triangularis, frontal orbital cortex, MFG) and right anterior insula had greater activation likelihood during higher demand than lower demand comprehension tasks in PSA (Fig. 5A, Supplementary Table S23).

Only 110 foci were obtained from 78 controls in 7 participant groups performing lower demand comprehension tasks. Accordingly, there were too few groups to perform ALE meta-analyses contrasting higher versus lower demand comprehension tasks in controls (Eickhoff et al., 2016). Thus, a third set of analyses combined PSA and control participant groups together. The single datasets from the meta-analysis comparing higher vs. lower demand comprehension tasks in PSA and control participants combined are reported in the Supplementary Information.

Contrast analyses revealed that a similar set of clusters in the left frontal lobe (frontal operculum, IFG pars opercularis/triangularis, MFG, precentral gyrus), left anterior insula, left temporooccipital ITG, midline cortex (SFG/SMC/paracingulate gyrus), right frontal lobe (frontal operculum, IFG pars opercularis/triangularis, frontal orbital cortex, MFG) and right anterior insula had greater activation likelihood during higher demand than lower demand comprehension tasks in PSA and controls combined (Fig. 5C, Supplementary Table S26).

These regions of increased activation likelihood during higher than lower demand comprehension tasks closely align with the semantic control network known to be involved during executively demanding semantic cognition in healthy individuals (Jackson, 2021) (Fig. 7B and C) and with the MD network (Fedorenko et al., 2013) (Fig. 6E, 6F). Critically, they overlap with clusters of greater activation likelihood in PSA than controls, across all language tasks and during comprehension tasks, in the right anterior insula and frontal operculum. They also overlap with PSA-specific clusters of increased activation likelihood during comprehension relative to production in the right anterior insula and MFG (Supplementary Table S18).

Contrast analyses revealed that activation was more likely in the left temporal pole in lower than higher demand comprehension tasks in PSA (Fig. 5A, Supplementary Table S23) and in both left and right temporal poles in lower than higher demand comprehension tasks in PSA and controls combined (Fig. 5 C, Supplementary Table S26).

3.3.2. Higher versus lower demand production tasks

Single datasets from the meta-analysis comparing higher vs. lower demand production tasks in PSA are reported in the Supplementary Information.

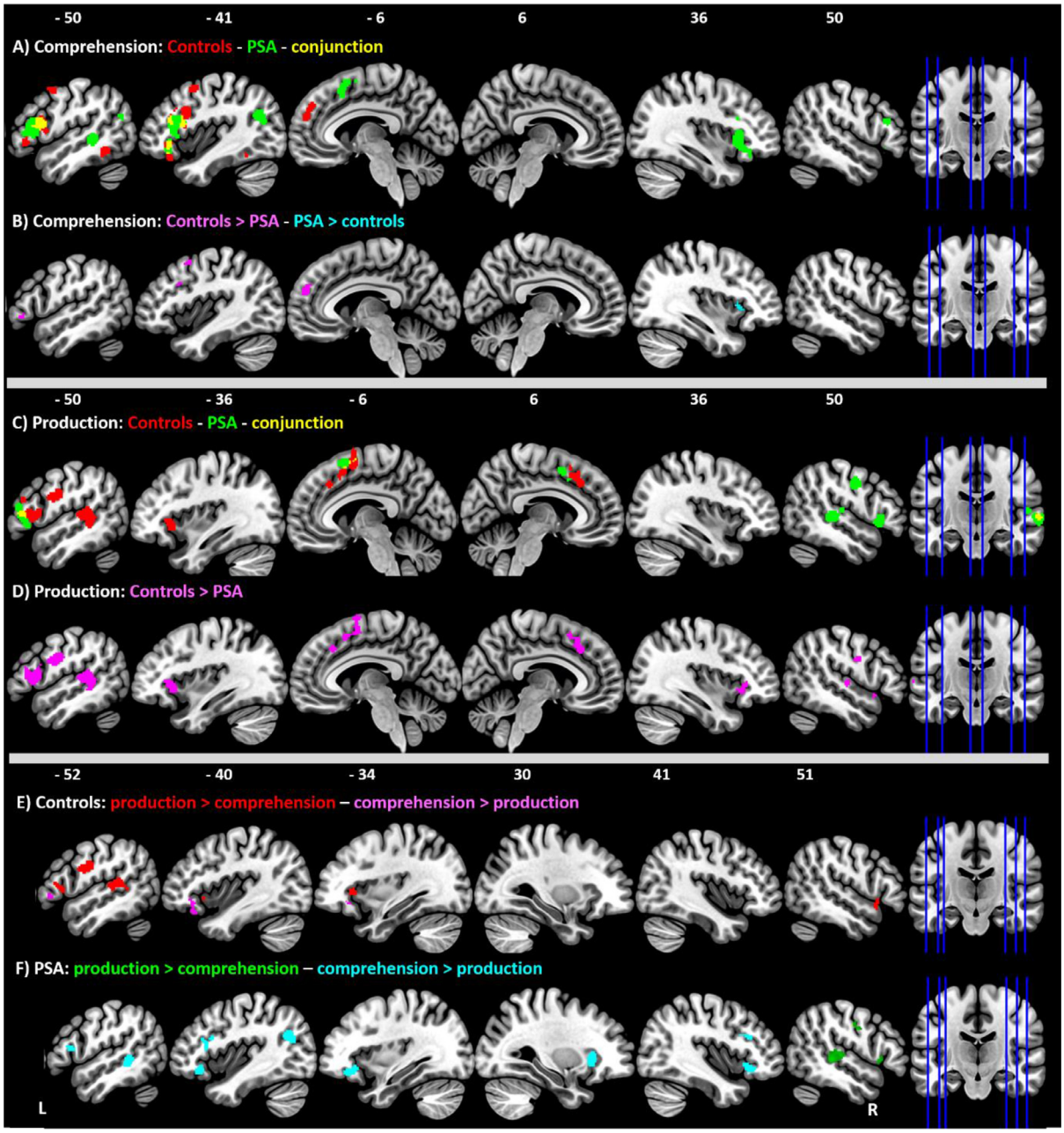


Fig. 4. ALE meta-analysis of comprehension and production tasks in PSA and healthy controls. A: ALE maps of comprehension tasks in PSA (green clusters) and in controls (red clusters), and conjunction map of comprehension tasks in both PSA and controls (yellow clusters). B: ALE maps of 'Comprehension tasks: controls > PSA' (violet clusters) and 'Comprehension tasks: PSA > controls' (cyan clusters). C: ALE maps of production tasks in PSA (green clusters), in controls (red clusters) and conjunction map of production tasks in both PSA and controls (yellow clusters). D: ALE maps of 'Production tasks: controls > PSA' (violet clusters). E: ALE maps of 'Controls: production > comprehension tasks' (red clusters), and 'Controls: comprehension > production tasks' (violet clusters). F: ALE maps of 'PSA: production > comprehension tasks' (green clusters), and 'PSA: comprehension > production tasks' (cyan clusters). Panels A and C: ALE single dataset analyses thresholded at $p<0.001$ uncorrected voxel-wise, FWE $p<0.05$ cluster wise, 1000 permutations. Panels B, D, E and F: ALE contrast analyses thresholded at $p<0.05$, 10000 permutations, minimum cluster extent 200 ml.

Contrast analyses revealed that the right frontal lobe (frontal operculum cortex, IFG pars opercularis/triangularis, precentral gyrus) and right temporal lobe (planum temporale, Heschl's gyrus) had greater activation likelihood during higher demand than lower demand production tasks in PSA (Fig. 5B, Supplementary Table S29).

Only 189 foci were obtained from 185 controls in 8 groups performing higher demand production tasks. Accordingly, there were too few groups to perform ALE meta-analyses contrasting higher versus lower demand production tasks in controls (Eickhoff et al., 2016). Thus, a third set of analyses combined PSA and control participant groups together. The single datasets from the meta-analysis comparing higher vs. lower demand production tasks in PSA and control participants combined are reported in the Supplementary Information.

Contrast analyses revealed that a similar set of clusters in the left IFG (frontal operculum, IFG pars opercularis/triangularis), left posterior MTG, right IFG (frontal operculum, IFG pars opercularis/triangularis, frontal orbital cortex), and right temporal lobe (Heschl's gyrus, planum temporale) had greater activation likelihood during higher demand than lower demand production tasks in PSA and controls combined (Fig. 5D, Supplementary Table S32).

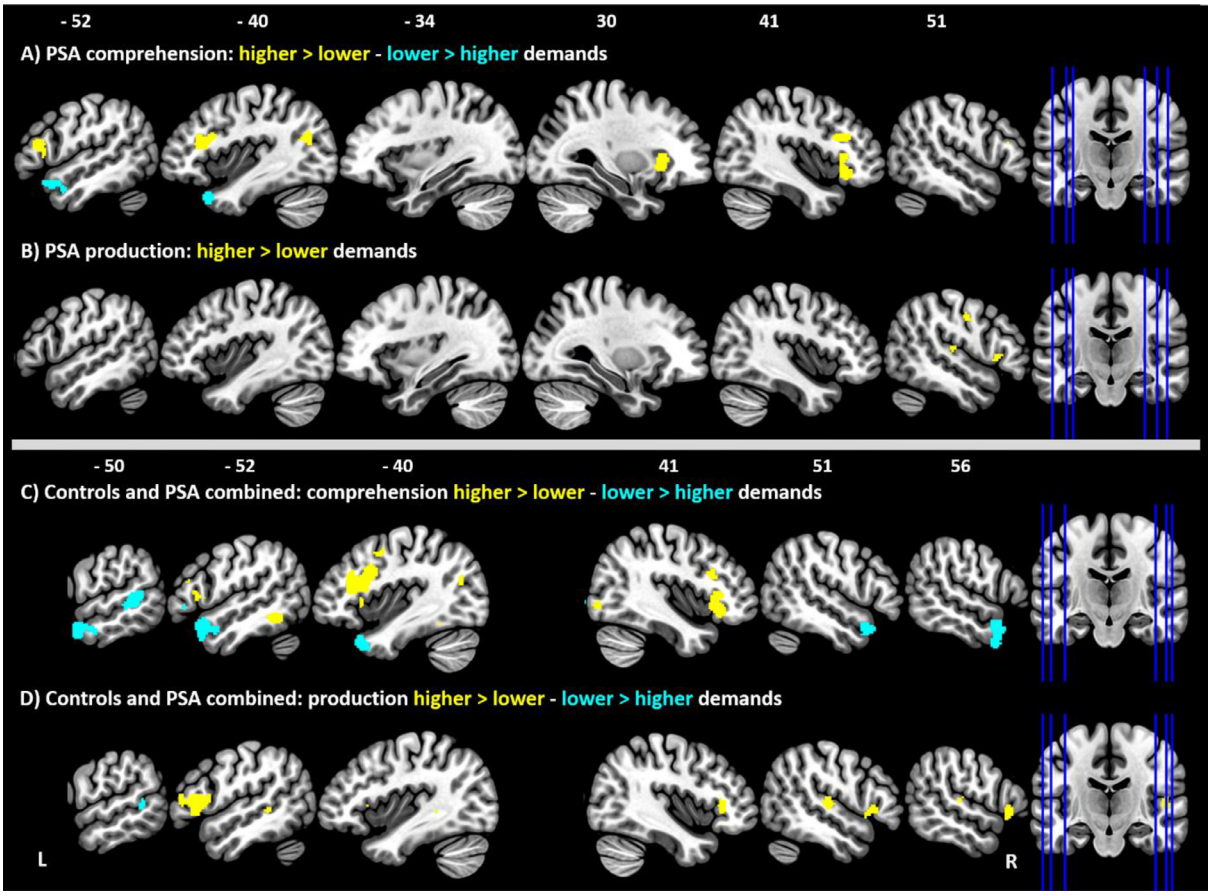


Fig. 5. Higher versus lower demands comprehension and production tasks. A: ALE maps of 'PSA comprehension tasks: higher > lower processing demands' (yellow clusters) and 'PSA comprehension tasks: lower > higher processing demands' (cyan clusters). B: ALE maps of 'PSA production tasks: higher > lower processing demands' (yellow clusters). C: ALE maps of 'PSA and healthy controls combined comprehension tasks: higher > lower processing demands' (yellow clusters) and 'PSA and healthy controls combined comprehension tasks: lower > higher processing demands' (cyan clusters). D: ALE maps of 'PSA and healthy controls combined production tasks: higher > lower processing demands' (yellow clusters) and 'PSA and healthy controls combined production tasks: lower > higher processing demands' (cyan clusters). All ALE contrast analyses thresholded at $p<0.05$, 10000 permutations, minimum cluster extent 200ml.

Right IFG clusters from the above difficulty-modulated production contrasts overlap with the MD network (Fedorenko et al., 2013) (Fig. 6E, 6F). Critically, they are also adjacent to clusters of greater activation likelihood in PSA than controls, across all language tasks and during comprehension tasks, in the right anterior insula and IFG. The right precentral gyrus difficulty-modulated production cluster in PSA alone overlapped with the PSA-specific cluster of increased activation likelihood during production relative to comprehension in the right precentral gyrus (Supplementary Table S18).

3.3.3. Summary

As predicted by *variable neurodisplacement*, right anterior insular and frontal opercular regions of greater activation likelihood in PSA than controls are more likely to be activated during more difficult than less difficult language tasks.

3.4. Time post-stroke

Our third aim was to investigate whether regions differentially activated in PSA relative to controls, vary between different stages of recovery. However, we found that the literature is strongly biased as most PSA underwent neuroimaging in the chronic phase post-stroke. The 64 PSA groups had median times post-stroke of 38.0 (IQR 34.5) months (Fig. 2). Only five papers, representing six of the 64 PSA groups, repeated functional neuroimaging longitudinally at multiple timepoints (Cardebat et al., 2003; Long et al., 2018; Nenert et al., 2018; Radman et al., 2016; Stockert et al., 2020). When counting the 'earliest' timepoint at which each PSA group was scanned, only 9/64 groups had mean times post-stroke less than 6 months (Cardebat et al., 2003; Geranmayeh et al., 2016; Long et al., 2018; Mattioli et al., 2014; Nenert et al., 2018; Qiu et al., 2017; Radman et al., 2016; Stockert et al., 2020). Accordingly, there were too few groups to contrast PSA before versus after six months (Eickhoff et al., 2016).

4. Discussion

In order to identify the specific regions that are more likely to be activated in PSA than healthy individuals, and to investigate whether there are differences in activation likelihood across different language tasks and between recovery timepoints, we performed a large-scale ALE meta-analysis of functional neuroimaging studies in PSA. We obtained coordinate-based functional neuroimaging data for 481 PSA, which is over four times larger than the last ALE meta-analysis on this topic (n=105) (Turkeltaub et al., 2011). The results provide novel insights into the mechanisms underlying language network changes post-stroke that might hitherto have been obscured by the limited sample size of any individual study in this area.

PSA were more likely to activate various regions of the right anterior insula and IFG than controls across all language tasks (anterior insula, frontal operculum, IFG pars opercularis) and during comprehension tasks (anterior insula, frontal operculum). These right anterior insular/IFG regions seem to be implicated in task difficulty as they are

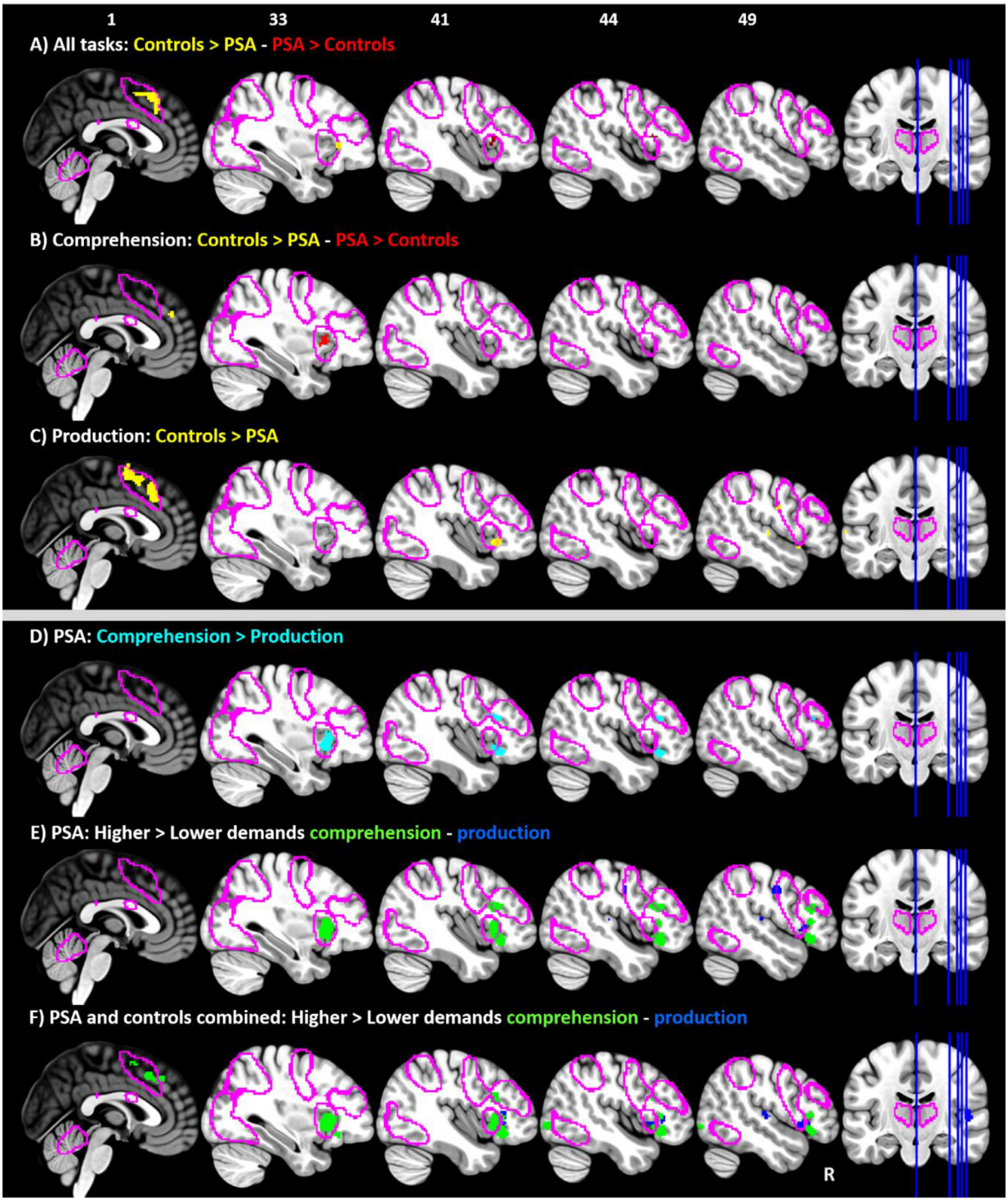


Fig. 6. Overlaps between clusters identified in the ALE meta-analyses and the Multiple Demand Network. A: ALE maps of 'Omnibus analysis: controls > PSA' (yellow clusters) and 'Omnibus analysis: PSA > controls' (red clusters). B: ALE maps of 'Comprehension: controls > PSA' (yellow clusters) and 'Comprehension: PSA > controls' (red clusters). C: ALE maps of 'Production: controls > PSA' (yellow clusters). D: ALE maps of 'Comprehension > production in PSA (cyan cluster). E: ALE maps of 'Comprehension higher > lower processing demands in PSA' (green cluster) and 'Production higher > lower processing demands in PSA' (blue cluster). F: ALE maps of 'Comprehension higher > lower processing demands in healthy controls and PSA combined' (green cluster) and 'Production higher > lower processing demands in healthy controls and PSA combined' (blue cluster). All panels include the outline of the Multiple Demand network (pink) (Fedorenko et al., 2013). All ALE contrast analyses thresholded at $p<0.05$, 10000 permutations, minimum cluster extent 200 ml.

more likely to be activated during higher than lower demand comprehension tasks (right anterior insula, frontal operculum, IFG pars opercularis/triangularis, frontal orbital cortex, in PSA and controls combined) and during higher than lower demand production tasks (right frontal operculum, IFG pars opercularis/triangularis, frontal orbital cortex, in PSA and controls combined). The networks subserving comprehension vs. production diverge in PSA relative to controls. Comprehension tasks in PSA make greater use of specific right frontal regions than both production tasks in PSA (right anterior insula, MFG), and comprehension tasks in controls (right anterior insula). Conversely, production tasks in

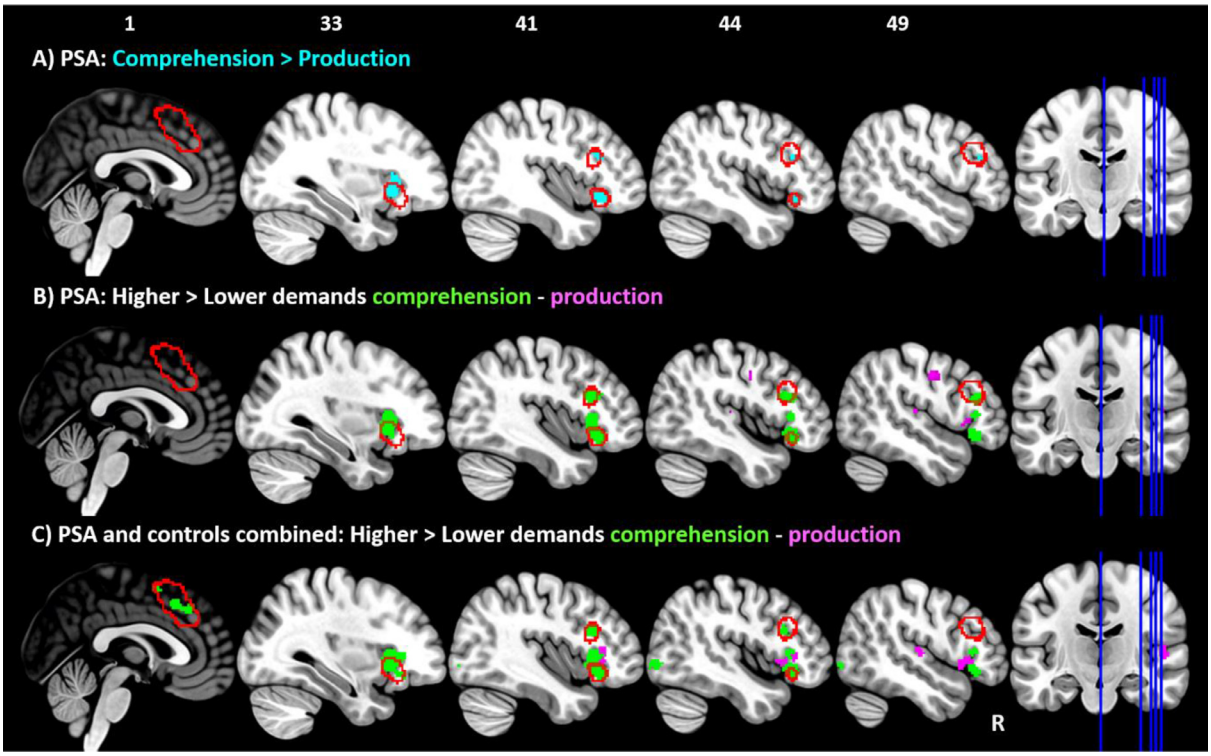


Fig. 7. Overlaps between clusters identified in the ALE meta-analyses and the Semantic Control Network. A: ALE maps of 'Comprehension > production in PSA (cyan cluster). B: ALE maps of 'Comprehension higher > lower processing demands in PSA' (green cluster) and 'Production higher > lower processing demands in PSA' (violet cluster). C: ALE maps of 'Comprehension higher > lower processing demands in healthy controls and PSA combined' (green cluster) and 'Production higher > lower processing demands in healthy controls and PSA combined' (violet cluster). All panels include the outline of the Semantic Control Network (red) (Jackson, 2021). All ALE contrast analyses thresholded at $p<0.05$, 10000 permutations, minimum cluster extent 200 ml.

PSA make greater use of a right precentral gyrus region than comprehension tasks in PSA, and this differential activation was not present in controls.

A previous ALE meta-analysis in PSA concluded that the language network in controls is left-lateralised, whereas PSA consistently activate additional homotopic right hemisphere regions that are not consistently activated in controls (Turkeltaub et al., 2011). The clear picture that emerges from the current, much larger ALE meta-analysis is different in a fundamental way. Whilst one can find reliably different levels of activation likelihood between the PSA and control groups, these differences all fall within regions that are found to activate in both groups; in classical neuropsychological terminology (Shallice, 1988), there is not a classical dissociation between PSA and control groups. Thus in the omnibus language ALE meta-analysis, the conjunction demonstrated that both PSA and controls consistently activated overlapping regions across the left and right frontal and temporal lobes, right parietal lobe, and midline cortex. Two important implications are that (a) right as well as left hemisphere areas make important contributions to language and (b) that regions, consistently activated by language tasks in PSA, are also involved in language pre-morbidly. This runs counter to the view that these areas are recruited 'de novo' post-stroke.

Irrespective of how the language tasks were divided (all language tasks, comprehension, production), we found that in PSA certain regions are less likely to be activated than in controls. These areas were not only left hemisphere regions that might have been lesioned directly by the stroke (i.e., within the left hemisphere MCA: cf. (Phan et al., 2005; Zhao et al., 2020)) but also domain-general regions of midline superior frontal and paracingulate cortex, right insula and right fronto-temporal cortex. This result implies that the language and cognitive deficits observed in PSA might not be a simple reflection of the lesioned areas but might result from combinations of lesioned and under-engaged areas. Accordingly, the use of task-based fMRI may be an important addition for future studies that aim to explore the neural bases of aphasia or build prediction models (Saur et al., 2010; Skipper-Kallal et al., 2017a; van Oers et al., 2018). Less consistent activation in regions distant to the lesions might reflect functional diaschisis, i.e., reduced task-related engagement throughout a connected network where one or more nodes have been compromised by damage (Carrera and Tononi, 2014). Alternatively from a more functional viewpoint, these distant regions may be less engaged because in PSA language is performed sub-optimally and therefore the full extent of the distributed language network is underutilised.

Neurocomputational invasion would predict that the post-stroke language network should expand to include novel non-language regions that were not consistently activated in healthy individuals (Keidel et al., 2010; Stefaniak et al., 2020). This mechanism is complementary to the classical notion that right hemisphere homologues of left hemisphere language regions are quiescent in health but become activated to perform similar language computations following left hemisphere stroke (Finger et al., 2003; Turkeltaub et al., 2011). A second linked idea is the notion of transcallosal disinhibition (Heiss and Thiel, 2006; Marshall, 1984). This proposes that right hemisphere, homologous regions are quiescent in health because they are inhibited transcallosally by the dominant left hemisphere, but can be 'released' when these dominant areas are damaged. This idea has been an important motivation for trials of non-invasive brain stimulation to inhibit the right IFG pars triangularis to aid language recovery through a shift back to left hemisphere areas (Bucur and Papagno, 2019; Ren et al., 2014). Previous work (Stefaniak et al., 2020) has noted that these hypotheses appear to be biologically-expensive (areas are maintained but not used, except in people who happen to have the right type and location of damage), computationally underspecified (e.g., how right hemisphere regions can

develop language functions when they are being constantly inhibited), and are an untested extension of findings from low-level, non-language motor circuitry (Di Lazzaro et al., 1999; Ferbert et al., 1992). Additional counter evidence includes: chronic language weaknesses can be found following right hemisphere damage (Gajardo-Vidal et al., 2018); and, residual language abilities in PSA have been related to the level of right hemisphere activation (Crinion and Price, 2005; Griffis et al., 2017; Skipper-Kallal et al., 2017b). The current study adds to these observations in that multiple regions throughout both hemispheres are consistently activated during language in both PSA and controls. Looking across these studies, it would seem that there is a solid empirical basis to move beyond oversimplified discussions of 'left versus right' language lateralisation and, instead, to explore how a bilateral, albeit asymmetrically left-biased, language network supports healthy function and generates aphasia after damage and partial recovery.

Variable neurodisplacement postulates that aphasia recovery involves increased utilisation of spare capacity within regions that are part of the premorbid language network but downregulated in health to save neural resources. Dynamic responses to performance demands in health and after damage could involve upregulation of language-specific and/or domain-general executive functions (Stefaniak et al., 2020). Accordingly, variable neurodisplacement encompasses the hypothesis that increased utilisation of domain-general executive regions aids language recovery post-stroke (Geranmayeh et al., 2014; Sharp et al., 2010). As noted above, a key finding from these ALE analyses was that bilateral regions, including domain-general parts of the MD network, were commonly engaged by PSA and control groups. Even where there were graded differences in favour of PSA over controls (e.g., greater activation likelihood in the right anterior insula and IFG), these are consistent with enhanced utilisation of demand-control regions due to increased task difficulty rather than 'expansion' into new territory via neurocomputational invasion. Thus, in the PSA group as well as PSA and controls combined, there was greater activation likelihood of the right anterior insula/operculum and IFG during higher than lower demand comprehension and production tasks. These same right anterior insula/IFG regions are known to be recruited during difficult tasks in healthy individuals: the right IFG has been implicated in domain-general top-down control in health (Baumgaertner et al., 2013; Koechlin and Jubault, 2006; Meinzer et al., 2012); a previous ALE meta-analysis found that effortful listening under difficult conditions in healthy individuals is associated with consistent activation in the bilateral insulae (Alain et al., 2018); and all ALE-identified right hemisphere regions overlap with either domain-general regions of the MD network (Fedorenko et al., 2013) or regions of the semantic control network known to be involved during executively-demanding semantic cognition in healthy individuals (Jackson, 2021).

The results do not suggest that there is a global, undifferentiated upregulation of all domain-general neural resources in PSA. Indeed, we repeatedly found lower activation likelihood in midline regions of the SFG/paracingulate gyrus in PSA compared to controls. These midline clusters overlap with at least some definitions of the domain-general executive network (Fedorenko et al., 2013). In contrast to our findings, increased activation in the same midline region has been associated with language recovery between two weeks and four months post-stroke (Geranmayeh et al., 2017). It is not clear what the basis of these opposing results is, but one possibility is that this ALE meta-analysis was predominantly based on data collected from patients in the very chronic (see below) rather than sub-acute stage. If correct, it may be the case that the executive functions supported by medial prefrontal regions (e.g., response conflict, task planning (Dosenbach et al., 2008; Mansouri et al., 2017)) are critical during early phases of recovery when performance is at its most impaired, but in relatively well-recovered, chronic PSA these mechanisms are not required (indeed continued involvement might signal poor recovery).

Activation was more likely in the anterior temporal lobes during lower than higher demand comprehension tasks. Previous task fMRI studies in healthy participants found minimal influence of semantic control demands in the anterior temporal lobe, unlike prefrontal or posterior temporal regions (Jackson, 2021). However, the anterior temporal lobe is more active for coherent, consistent contexts and combinatorial meanings, while inconsistent context or combinations of meaning require increased activation in semantic control and executive demand areas (Branzi et al., 2020; Hoffman et al., 2015). Consequently, the anterior temporal lobe result in the current ALE meta-analysis might reflect comprehension processes for coherent contexts and combinations during lower demand comprehension tasks.

As is commonly the case in stroke research (Fareed et al., 2012; Thomalla et al., 2017), the median ages of the 64 included PSA participant groups was lower (57.4 years) than the average stroke patient (e.g., the median age of the UK stroke population was 77 in 2017 (SSNAP, 2017). This may limit the generalisability of results obtained from functional neuroimaging studies to the 'real-world', and future studies should investigate patterns of activation in older PSA that are more representative of the average stroke survivor.

We identified areas of enquiry that have had little attention in the literature to date. It was not possible to ascertain whether there are consistent activation differences between subacute and chronic PSA. The 64 PSA groups had median times post-stroke of 38.0 months and even when counting the 'earliest' timepoint at which each PSA group was scanned, only 9/64 PSA groups were less than 6 months post-stroke. This dearth of data meant it was not possible to use ALE to explore differences between sub-acute and chronic PSA. Importantly, this indicates a pressing need for future studies of this early period, when there is the fastest rate of language recovery (Pedersen et al., 1995; Yagata et al., 2017). Additionally, it was not possible to explore longitudinal fMRI changes given the extremely limited number of longitudinal PSA fMRI studies. Even among papers that reported longitudinal information, several were small (n<10 participants) and there was considerable variation with respect to which language or non-language cognition was explored and the timing of the first imaging timepoint (from the first few days to a few months post-stroke). The relative lack of studies and small sample sizes are unsurprising given the considerable logistic challenges involved in imaging subacute stroke patients. However, longitudinal studies are a powerful approach for exploring the neural bases of recovery (because the different starting points and inter-participant variations are controlled), and particularly for exploring whether language network changes observed in the chronic phase occur immediately or over time. Such information will be critical for understanding the mechanisms underpinning both instantaneous resilience to the effects of damage, degeneracy, and longer-term experience-dependent plasticity (Chang and Lambon Ralph, 2020; Price and Friston, 2002; Sajid et al., 2020; Stefaniak et al., 2020; Ueno et al., 2011).

The results of this large-scale meta-analysis argue against classical neurocomputational invasion accounts of PSA language, i.e., expansion of the language network post-damage into new territories. Instead, (a) there is considerable overlap between the bilateral language-related functional networks observed in PSA and controls; (b) the PSA participants are less likely than controls to activate certain regions including areas beyond their core lesions in the left MCA territory; and (c) are more likely to engage executive-control related regions of the right anterior insula and IFG. These results fit with a view that language is supported by a dynamic, bilateral albeit left-asymmetric network, and consistent with the variable neurodisplacement hypothesis. The size of this (random-effects) analysis (including data pertaining to 481 PSA with a heterogenous variety of lesion locations and aphasia profiles), should mean that the results will generalise to the wider patient population.

Despite its size and clear results, inevitably this study has limitations. First, all included PSA participants had a single left hemisphere stroke, so it is possible that left hemisphere clusters of lower activation likelihood in PSA might be a direct effect of tissue damage. Relatedly, left hemisphere lesions might have biased single dataset meta-analyses of PSA participants towards consistency in the right hemisphere, al-

though there would have been no right hemisphere biasing effect on single dataset meta-analyses of controls, nor on any of the contrast meta-analyses. Second, decreased neurovascular coupling post-stroke could generate false activation differences between patients and controls, although this is less likely in chronic patients and undamaged cortical regions (Geranmayeh et al., 2015). Third, 'neural reprogramming' might entail differences in utilisation that are only observable using connectivity (Meier et al., 2018; Schofield et al., 2012) or multivariate analyses (Fischer-Baum et al., 2017; Lee et al., 2017), although very few studies have used such techniques in PSA to date. Finally, the meta-analysis rests on studies reporting the full set of whole brain responses from both PSA and controls, and differences seen in meta-analyses might not be replicated in individual studies.

Declaration of Competing Interests

None.

Author Contributions

JDS and RSWA contributed to study design, data collection, analysis and write-up. MALR contributed to study design and write-up.

Credit Author Statement

JDS: Conceptualization, Formal analysis, Investigation, Data curation, Writing. RSWA: Conceptualization, Formal analysis, Investigation, Data curation, Writing, Visualization. MALR: Conceptualization, Writing, Supervision.

Data availability

Group level coordinate data supporting the findings of this study are available on figshare (doi: 10.6084/m9.figshare.12582935).

Acknowledgements

We would like to thank Sonia Brownsett, Fatemeh Geranmayeh, Glyn Hallam, Jessica Hodgson, Narges Radman, Michael Mouthon, Alex Leff, Stephen Wilson, Yuan Tao, Erin Meier, and Elena Barbieri for kindly sharing unpublished coordinates relating to their published studies for use in this ALE meta-analysis. JDS was a Wellcome clinical PhD fellow funded on grant 203914/Z/16/Z to the Universities of Manchester, Leeds, Newcastle and Sheffield. MALR is supported by an European Research Council Advanced grant (GAP: 670428), a Rosetrees Trust grant (A1699) and MRC intra-mural funding (MC_UU_00005/18). The funding sources had no involvement in the study.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.117960.

References

Alain, C., Du, Y., Bernstein, L.J., Barten, T., Banai, K., 2018. Listening under difficult conditions: an activation likelihood estimation meta-analysis. Hum. Brain Mapp. 39, 2695–2709.

Alyahya, R.S.W., Halai, A., Conroy, P., Lambon Ralph, M.A., 2020. A unified model of post-stroke language deficits including discourse production and their neural correlates. Brain 143, 1541–1554.

Barbieri, E., Mack, J., Chiappetta, B., Europa, E., Thompson, C.K., 2019. Recovery of offline and online sentence processing in aphasia: language and domain-general network neuroplasticity. Cortex 120, 394–418.

Baumgaertner, A., Hartwigsen, G., Roman Siebner, H., 2013. Right-hemispheric processing of non-linguistic word features: implications for mapping language recovery after stroke. Hum. Brain Mapp. 34, 1293–1305.

Binney, R.J., Lambon Ralph, M.A., 2015. Using a combination of fMRI and anterior temporal lobe rTMS to measure intrinsic and induced activation changes across the semantic cognition network. Neuropsychologia 76, 170–181.

Branzi, F.M., Humphreys, G.F., Hoffman, P., Lambon Ralph, M.A., 2020. Revealing the neural networks that extract conceptual gestalts from continuously evolving or changing semantic contexts. Neuroimage 220, 116802.

Bucur, M., Papagno, C., 2019. Are transcranial brain stimulation effects long-lasting in post-stroke aphasia? A comparative systematic review and meta-analysis on naming performance. Neurosci. Biobehav. Rev. 102, 264–289.

Butler, R.A., Lambon Ralph, M.A., Woollams, A.M., 2014. Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. Brain 137, 3248–3266.

Cardebat, D., Demonet, J.F., De Boissezon, X., Marie, N., Marie, R.M., Lambert, J., Baron, J.C., Puel, M., Study, P.E.T.L.A., 2003. Behavioral and neurofunctional changes over time in healthy and aphasic subjects: a PET language activation study. Stroke 34, 2900–2906.

Carrera, E., Tononi, G., 2014. Diaschisis: past, present, future. Brain 137, 2408–2422.

Chang, Y.-N., Lambon Ralph, M., 2020. A unified neurocomputational bilateral pathway model of spoken language production in healthy participants and recovery in post-stroke aphasia. Proc. Natl. Acad. Sci. USA 117, 32779–32790.

Crinion, J., Price, C.J., 2005. Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. Brain 128, 2858–2871.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968–980.

Di Lazzaro, V., Oliviero, A., Profice, P., Insola, A., Mazzone, P., Tonali, P., Rothwell, J., 1999. Direct demonstration of interhemispheric inhibition of the human motor cortex produced by transcranial magnetic stimulation. Exp. Brain Res. 124, 520–524.

Dosenbach, N.U., Fair, D.A., Cohen, A.L., Schlaggar, B.L., Petersen, S.E., 2008. A dual-networks architecture of top-down control. Trends Cogn. Sci. 12, 99–105.

Duncan, J., 2010. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends Cogn. Sci. 14, 172–179.

Eickhoff, S., Bzdok, D., 2013. Meta-analyses in basic and clinical neuroscience: state of the art and perspective. In: S, U., O, J. (Eds.), fMRI. Springer, Berlin, Heidelberg, pp. 77–87.

Eickhoff, S.B., Bzdok, D., Laird, A.R., Kurth, F., Fox, P.T., 2012. Activation likelihood estimation meta-analysis revisited. Neuroimage 59, 2349–2361.

Eickhoff, S.B., Bzdok, D., Laird, A.R., Roski, C., Caspers, S., Zilles, K., Fox, P.T., 2011. Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. Neuroimage 57, 938–949.

Eickhoff, S.B., Laird, A.R., Grefkes, C., Wang, L.E., Zilles, K., Fox, P.T., 2009. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. Hum. Brain Mapp. 30, 2907–2926.

Eickhoff, S.B., Nichols, T.E., Laird, A.R., Hoffstaedter, F., Amunts, K., Fox, P.T., Bzdok, D., Eickhoff, C.R., 2016. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. Neuroimage 137, 70–85.

Engelter, S.T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., Gutzwiller, F., Lyrer, P.A., 2006. Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. Stroke 37, 1379–1384.

Fareed, M., Suri, K., Qureshi, A., 2012. Recruitment of ischemic stroke patients in clinical trials in general practice and implications for generalizability of results. J. Vasc. Interv. Neurol. 5, 27–32.

Fedorenko, E., Behr, M.K., Kanwisher, N., 2011. Functional specificity for high-level linguistic processing in the human brain. Proc. Natl. Acad. Sci. USA 108, 16428–16433.

Fedorenko, E., Duncan, J., Kanwisher, N., 2013. Broad domain generality in focal regions of frontal and parietal cortex. Proc. Natl. Acad. Sci. USA 110, 16616–16621.

Ferbert, A., Priori, A., Rothwell, J.C., Day, B.L., Colebatch, J.G., Marsden, C.D., 1992. Interhemispheric inhibition of the human motor cortex. J. Physiol. 453, 525–546.

Finger, S., Buckner, R.L., Buckingham, H., 2003. Does the right hemisphere take over after damage to Broca's area? the Barlow case of 1877 and its history. Brain Lang. 85, 385–395.

Fischer-Baum, S., Jang, A., Kajander, D., 2017. The cognitive neuroplasticity of reading recovery following chronic stroke: a representational similarity analysis approach. Neural Plasticity, 2761913.

Gajardo-Vidal, A., Lorca-Puls, D.L., Hope, T.M.H., Parker Jones, O., Seghier, M.L., Prejawa, S., Crinion, J.T., Leff, A.P., Green, D.W., Price, C.J., 2018. How right hemisphere damage after stroke can impair speech comprehension. Brain 141, 3389–3404.

Geranmayeh, F., Brownsett, S.L., Wise, R.J., 2014. Task-induced brain activity in aphasic stroke patients: what is driving recovery? Brain 137, 2632–2648.

Geranmayeh, F., Chau, T.W., Wise, R.J.S., Leech, R., Hampshire, A., 2017. Domain-general subregions of the medial prefrontal cortex contribute to recovery of language after stroke. Brain 140, 1947–1958.

Geranmayeh, F., Leech, R., Wise, R.J., 2016. Network dysfunction predicts speech production after left hemisphere stroke. Neurology 86, 1296–1305.

Geranmayeh, F., Wise, R.J., Leech, R., Murphy, K., 2015. Measuring vascular reactivity with breath-holds after stroke: a method to aid interpretation of group-level BOLD signal changes in longitudinal fMRI studies. Hum. Brain Mapp. 36, 1755–1771.

Gordon, P.C., Hendrick, R., Levine, W.H., 2002. Memory-load interference in syntactic processing. Psychol. Sci. 13, 425–430.

Griffis, J.C., Nenert, R., Allendorfer, J.B., Vannest, J., Holland, S., Dietz, A., Szaflarski, J.P., 2017. The canonical semantic network supports residual language function in chronic post-stroke aphasia. Hum. Brain Mapp. 38, 1636–1658.

Hallam, G.P., Thompson, H.E., Hymers, M., Millman, R.E., Rodd, J.M., Lambon Ralph, M.A., Smallwood, J., Jefferies, E., 2018. Task-based and resting-state fMRI reveal compensatory network changes following damage to left inferior frontal gyrus. Cortex 99, 150–165.

Hartwigsen, G., 2018. Flexible redistribution in cognitive networks. Trends Cogn. Sci. 22, 687–698.
Heiss, W.D., Thiel, A., 2006. A proposed regional hierarchy in recovery of post-stroke aphasia. Brain Lang. 98, 118–123.
Hoffman, P., Binney, R.J., Lambon Ralph, M.A., 2015. Differing contributions of inferior prefrontal and anterior temporal cortex to concrete and abstract conceptual knowledge. Cortex 63, 250–266.
Hope, T.M.H., Leff, A.P., Prejawa, S., Bruce, R., Haigh, Z., Lim, L., Ramsden, S., Oberhuber, M., Ludersdorfer, P., Crinion, J., Seghier, M.L., Price, C.J., 2017. Right hemisphere structural adaptation and changing language skills years after left hemisphere stroke. Brain 140, 1718–1728.
Jackson, R.L., 2021. The neural correlates of semantic control revisited. Neuroimage 224, 117444.
Jung, J., Lambon Ralph, M.A., 2016. Mapping the dynamic network interactions underpinning cognition: a cTBS-fMRI study of the flexible adaptive neural system for semantics. Cerebr. Cortex 26, 3580–3590.
Keidel, J.L., Welbourne, S.R., Lambon Ralph, M.A., 2010. Solving the paradox of the equipotential and modular brain: a neurocomputational model of stroke vs. slow–growing glioma. Neuropsychologia 48, 1716–1724.
Koechlin, E., Jubault, T., 2006. Broca's area and the hierarchical organization of human behavior. Neuron 50, 963–974.
Kummerer, D., Hartwigsen, G., Kellmeyer, P., Glauche, V., Mader, I., Kloppel, S., Suchan, J., Karnath, H.O., Weiller, C., Saur, D., 2013. Damage to ventral and dorsal language pathways in acute aphasia. Brain 136, 619–629.
Lambon Ralph, M.A., McClelland, J.L., Patterson, K., Galton, C.J., Hodges, J.R., 2001. No right to speak? The relationship between object naming and semantic impairment: neuropsychological evidence and a computational model. J. Cogn. Neurosci. 13, 341–356.
Lancaster, J.L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., Mazziotta, J.C., Fox, P.T., 2007. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum. Brain Mapp. 28, 1194–1205.
Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T., 2000. Automated Talairach atlas labels for functional brain mapping. Hum. Brain Mapp. 10, 120–131.
Lee, Y.S., Zreik, J.T., Hamilton, R.H., 2017. Patterns of neural activity predict picture–naming performance of a patient with chronic aphasia. Neuropsychologia 94, 52–60.
Lidzba, K., Schwilling, E., Grodd, W., Krägeloh-Mann, I., Wilke, M., 2011. Language comprehension vs. language production: age effects on fMRI activation. Brain Lang. 119, 6–15.
Long, C., Sebastian, R., Faria, A.V., Hillis, A.E., 2018. Longitudinal imaging of reading and naming recovery after stroke. Aphasiology 32, 839–854.
Mansouri, F., Egner, T., Buckley, M., 2017. Monitoring demands for executive control: shared functions between human and nonhuman primates. Trends Neurosci. 40, 15–27.
Marshall, J.F., 1984. Brain function: neural adaptations and recovery from injury. Annu. Rev. Psychol. 35, 277–308.
Mattioli, F., Ambrosi, C., Mascaro, L., Scarpazza, C., Pasquali, P., Frugoni, M., Magoni, M., Biagi, L., Gasparotti, R., 2014. Early aphasia rehabilitation is associated with functional reactivation of the left inferior frontal gyrus: a pilot study. Stroke 45, 545–552.
Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Perchey, G., Mellet, E., Petit, L., Tzourio-Mazoyer, N., 2014. Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. PLoS One 9, e101165.
Meier, E.L., Johnson, J.P., Kiran, S., 2018. Left frontotemporal effective connectivity during semantic feature judgments in patients with chronic aphasia and age-matched healthy controls. Cortex 108, 173–192.
Meier, E.L., Johnson, J.P., Pan, Y., Kiran, S., 2019. A lesion and connectivity-based hierarchical model of chronic aphasia recovery dissociates patients and healthy controls. NeuroImage Clin. 23, 101919.
Meinzer, M., Flaisch, T., Seeds, L., Harnish, S., Antonenko, D., Witte, V., Lindenberg, R., Crosson, B., 2012. Same modulation but different starting points: Performance modulates age differences in inferior frontal cortex activity during word-retrieval. PLoS One 7, e33631.
Mementi, L., Gierhan, S., Segaert, K., Hagoort, P., 2011. Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. Psychol. Sci. 22, 1173–1182.
Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O.K., Coslett, H.B., Schwartz, M.F., 2015. Neural organization of spoken language revealed by lesion-symptom mapping. Nat. Commun. 6, 6762.
Murphy, T.H., Corbett, D., 2009. Plasticity during stroke recovery: from synapse to behaviour. Nat. Rev. Neurosci. 10, 861–872.
Müller, V.I., Cieslik, E.C., Laird, A.R., Fox, P.T., Radua, J., Mataix-Cols, D., Tench, C.R., Yarkoni, T., Nichols, T.E., Turkeltaub, P.E., Wager, T.D., Eickhoff, S.B., 2018. Ten simple rules for neuroimaging meta-analysis. Neurosci. Biobehav. Rev. 84, 151–161.
Nenert, R., Allendorfer, J.B., Martin, A.M., Banks, C., Vannest, J., Holland, S.K., Hart, K.W., Lindsell, C.J., Szaflarski, J.P., 2018. Longitudinal fMRI study of language recovery after a left hemispheric ischemic stroke. Restor. Neurol. Neurosci. 36, 359–385.
Patterson, K., Lambon Ralph, M.A., 1999. Selective disorders of reading? Curr. Opin. Neurobiol. 9, 235–239.
Pedersen, P.M., Jørgensen, H.S., Nakayama, H., Raaschou, H.O., Olsen, T.S., 1995. Aphasia in acute stroke: incidence, determinants, and recovery. Ann. Neurol. 38, 659–666.
Phan, T., Donnan, G., Wright, P., Reutens, D., 2005. A digital map of middle cerebral artery infarcts associated with middle cerebral artery trunk and branch occlusion. Stroke 36, 986–991.
Price, C.J., Friston, K.J., 2002. Degeneracy and cognitive anatomy. Trends Cogn. Sci. 6, 416–421.
Pritchett, B.L., Hoeflin, C., Koldewyn, K., Dechter, E., Fedorenko, E., 2018. High-level language processing regions are not engaged in action observation or imitation. J. Neurophysiol. 120, 2555–2570.
Qiu, W.H., Wu, H.X., Yang, Q.L., Kang, Z., Chen, Z.C., Li, K., Qiu, G.R., Xie, C.Q., Wan, G.F., Chen, S.Q., 2017. Evidence of cortical reorganization of language networks after stroke with subacute Broca's aphasia: a blood oxygenation level dependent-functional magnetic resonance imaging study. Neural Regen. Res. 12, 109–117.
Radman, N., Mouthon, M., Di Pietro, M., Gaytanidis, C., Leemann, B., Abutalebi, J., Annoni, J.M., 2016. The role of the cognitive control system in recovery from bilingual aphasia: a multiple single-case fMRI study. Neural Plast. 8797086.
Ren, C.L., Zhang, G.F., Xia, N., Jin, C.H., Zhang, X.H., Hao, J.F., Guan, H.B., Tang, H., Li, J.A., Cai, D.L., 2014. Effect of low-frequency rTMS on aphasia in stroke patients: a meta-analysis of randomized controlled trials. PLoS One 9, e102557.
Rice, G.E., Caswell, H., Moore, P., Lambon Ralph, M.A., Hoffman, P., 2018. Revealing the dynamic modulations that underpin a resilient neural network for semantic cognition: an fMRI investigation in patients with anterior temporal lobe resection. Cereb Cortex 28, 3004–3016.
Robson, H., Zahn, R., Keidel, J.L., Binney, R.J., Sage, K., Lambon Ralph, M.A., 2014. The anterior temporal lobes support residual comprehension in Wernicke's aphasia. Brain 137, 931–943.
Sajid, N., Parr, T., Hope, T., Price, C., Friston, K., 2020. Degeneracy and redundancy in active inference. Cereb. Cortex 30, 5750–5766.
Saur, D., Ronneberger, O., Kummerer, D., Mader, I., Weiller, C., Kloppel, S., 2010. Early functional magnetic resonance imaging activations predict language outcome after stroke. Brain 133, 1252–1264.
Schofield, T.M., Penny, W.D., Stephan, K.E., Crinion, J.T., Thompson, A.J., Price, C.J., Leff, A.P., 2012. Changes in auditory feedback connections determine the severity of speech processing deficits after stroke. J. Neurosci. 32, 4260–4270.
Shallice, T., 1988. From Neuropsychology to Mental Structure. Cambridge University Press, New York, NY.
Sharp, D.J., Turkheimer, F.E., Bose, S.K., Scott, S.K., Wise, R.J., 2010. Increased frontoparietal integration after stroke and cognitive recovery. Ann. Neurol. 68, 753–756.
Skipper-Kallal, L.M., Lacey, E.H., Xing, S., Turkeltaub, P.E., 2017a. Functional activation independently contributes to naming ability and relates to lesion site in post-stroke aphasia. Hum. Brain Mapp. 38, 2051–2066.
Skipper-Kallal, L.M., Lacey, E.H., Xing, S., Turkeltaub, P.E., 2017b. Right hemisphere remapping of naming functions depends on lesion size and location in poststroke aphasia. Neural Plast. 8740353.
Southwell, D.G., Hervey-Jumper, S.L., Perry, D.W., Berger, M.S., 2016. Intraoperative mapping during repeat awake craniotomy reveals the functional plasticity of adult cortex. J. Neurosurg. 124, 1460–1469.
SSNAP, 2017. Sentinel Stroke National Audit Programme (SSNAP).
Stefaniak, J.D., Halai, A.D., Lambon Ralph, M.A., 2020. The neural and neurocomputational bases of recovery from post-stroke aphasia. Nat. Rev. Neurol. 16, 43–55.
Stockert, A., Wawrzyniak, M., Klingbeil, J., Wrede, K., Kummerer, D., Hartwigsen, G., Kaller, C.P., Weiller, C., Saur, D., 2020. Dynamics of language reorganization after left temporo-parietal and frontal stroke. Brain: J. Neurol. 143, 844–861.
Tao, Y., Rapp, B., 2019. The effects of lesion and treatment-related recovery on functional network modularity in post-stroke dysgraphia. NeuroImage Clin. 23, 101865.
Thomalla, G., Boutitie, F., Fiebach, J.B., Simonsen, C.Z., Nighoghossian, N., Pedraza, S., Lemmens, R., Roy, P., Muir, K.W., Heesen, C., Ebinger, M., Ford, I., Cheng, B., Cho, T.H., Puig, J., Thijs, V., Endres, M., Fiehler, J., Gerloff, C., 2017. Effect of informed consent on patient characteristics in a stroke thrombolysis trial. Neurology 89, 1400–1407.
Turkeltaub, P.E., Eickhoff, S.B., Laird, A.R., Fox, M., Wiener, M., Fox, P., 2012. Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. Hum. Brain Mapp. 33, 1–13.
Turkeltaub, P.E., Messing, S., Norise, C., Hamilton, R.H., 2011. Are networks for residual language function and recovery consistent across aphasic patients? Neurology 76, 1726–1734.
Ueno, T., Saito, S., Rogers, T., Lambon Ralph, M., 2011. Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. Neuron 72, 385–396.
van Oers, C., van der Worp, H.B., Kappelle, L.J., Raemaekers, M.A.H., Otte, W.M., Dijkhuizen, R.M., 2018. Etiology of language network changes during recovery of aphasia after stroke. Sci. Rep. 8, 856.
Warburton, E., Price, C.J., Swinburn, K., Wise, R.J., 1999. Mechanisms of recovery from aphasia: evidence from positron emission tomography studies. J. Neurol., Neurosurg. Psychiatry 66, 155–161.
Wilson, S.M., Yen, M., Eriksson, D.K., 2018. An adaptive semantic matching paradigm for reliable and valid language mapping in individuals with aphasia. Hum. Brain Mapp. 39, 3285–3307.
Yagata, S.A., Yen, M., McCarron, A., Bautista, A., Lamair-Orosco, G., Wilson, S.M., 2017. Rapid recovery from aphasia after infarction of Wernicke's area. Aphasiology 31, 951–980.
Zhao, Y., Halai, A., Lambon Ralph, M., 2020. Evaluating the granularity and statistical structure of lesions and behaviour in post-stroke aphasia. Brain Commun. 2, fcaa062.

Chapter 4

Information diaschisis within language and domain-general regions in chronic post-stroke aphasia

Manuscript prepared for submission

Abstract

Aphasia frequently persists into the chronic phase post-stroke, but our understanding of the mechanisms underlying language recovery is limited. A central research aim is to understand which neural regions become more functionally involved in language to promote recovery, and whether such regions are language-specific or domain-general. Previous studies have focused almost exclusively on univariate language activation differences between post-stroke aphasia and controls, but information-based multivariate pattern analysis might be more apposite for identifying regions functionally involved in language. Accordingly, this study: performed a language multivariate pattern analysis experiment in post-stroke aphasia and controls; identified regions in which language information content associated positively with language performance; and assessed whether this relationship was specific to language decoding or could also be observed using univariate language or non-language executive activation. Twenty-four patients with aphasia at least 6 months post left-hemispheric stroke and 30 controls underwent functional magnetic resonance imaging and behavioural testing. The primary language task consisted of passively listening to sentences with high versus low semantic predictability endings; a second scanner task consisted of hard versus easy visuospatial pattern matching to identify 'domain-general cognitive difficulty' regions. Clusters of significant sentence cloze decoding were present in core regions of the bilateral semantic control network but additionally included regions that were not activated (left supramarginal gyrus) or even deactivated (precuneus) during the univariate language contrast, and might therefore have been overlooked were activation considered in isolation. Language information content was significantly more positively associated with language ability throughout bilateral fronto-temporo-parietal cortex in participants with aphasia than controls. Critically, these regions had decreased decoding accuracy in participants with aphasia, suggesting that they represented a novel form of 'information diaschisis' in which a stroke triggers lower language information processing in distant undamaged nodes of the residual language network which, in turn, contributes to the language impairment. Univariate language activation in these regions was not associated with language performance, showing that pattern decoding might be useful as a novel biomarker and that information content may be more important than activation during aphasia recovery. Most regions in which language information content correlated with language performance were domain-general, suggesting that domain-general executive regions process and/or detect high-level language information in health, even when executive demands are minimal. Together, these results suggest that it is specifically the maintenance of language information processing, rather than univariate language

or non-language activation, in a bilateral set of predominantly domain-general regions that is important for retention and recovery of language performance post-stroke.

Introduction

Aphasia is a highly prevalent (Engelter et al., 2006) cause of morbidity (Ellis et al., 2012) post-stroke that is associated with increased care costs (Boehme et al., 2016), functional dependence and death (Tsouli et al., 2009). Unfortunately, language recovery can be incomplete and aphasia often persists into the chronic phase (Maas et al., 2012; Pedersen et al., 1995; Wade et al., 1986). Despite considerable research efforts, our understanding of the mechanisms underlying aphasia recovery is very limited (Stefaniak et al., 2020). An increasing number of functional neuroimaging studies have therefore sought to identify the neural basis of aphasia recovery by investigating whether different neural regions are functionally involved in language in post-stroke aphasia (PSA) relative to controls (Stefaniak et al., 2021). If the functional involvement of a region is found to be positively associated with language performance in PSA and it is more functionally involved in PSA than controls, it might serve a compensatory role to aid aphasia recovery. Conversely, if that region is less functionally involved in PSA than controls, it might represent 'functional diaschisis', in which reduced engagement of distant undamaged nodes of a connected network contributes to the language impairment (Carrera & Tononi, 2014; Stefaniak et al., 2021). Such investigations might illuminate which of several proposed recovery mechanisms occur in vivo, provide neurobiologically-informed therapeutic targets for non-invasive brain stimulation, or enable the development of neuroimaging-based biomarkers to predict clinical outcome (Stefaniak et al., 2020). Nevertheless, uncertainty remains as to which neural regions become more or less functionally involved in language post-stroke.

A striking feature of the current literature is that functional involvement during language has been defined almost exclusively in terms of activation. Activation-based functional brain imaging identifies regions whose activity changes between two task conditions. The resultant activation differences are used to infer which regions might be relevant to that task. The most commonly used analyses are mass-univariate techniques, which consider the activation of each voxel separately, although more recent multivariate techniques such as independent component analysis (McKeown et al., 2003) have identified task-related networks of co-varying voxels (Geranmayeh et al., 2016). In contrast to these activation-based approaches, more recent neuroimaging explorations of healthy performance have developed ‘information-based’ approaches (Kriegeskorte et al., 2006), collectively termed ‘multivariate pattern analysis’

(MVPA) (Haynes, 2015), that assess whether the distributed activity pattern across multiple voxels in a region can be used to predict (or 'decode') which of two cognitive states a participant is in.

MVPA could offer several advantages for investigating the neural bases of PSA recovery. First, MVPA is more sensitive than univariate activation approaches at identifying regions involved in a wide range of cognitive tasks (Haxby et al., 2001; Haynes et al., 2007). This is because univariate activation approaches assume all voxels respond in the same way to a cognitive task (with increased or decreased activity). However, higher cognitive functions such as language involve distributed, overlapping representations (Carota et al., 2017; Haxby et al., 2001) in which different voxels might process different dimensions of a cognitive task. When analysed separately, the activation change of individual voxels might be of insufficient magnitude to be distinguishable from noise. Using MVPA, these 'distributed multidimensional effects' might convey sufficient information to be detected (Davis et al., 2014). It is therefore plausible that MVPA might identify neural correlates of language recovery better than univariate activation imaging in PSA, just as MVPA classification accuracy correlates better with symptom severity than mean activation in autism (Coutanche et al., 2011). However, no published study has yet investigated whether decoding accuracy is associated with out-of-scanner language performance in PSA. Secondly, activation by itself is probably insufficient to contribute to the performance of a cognitive task; rather, activation must help process task-relevant behavioural information and internal representations to contribute to behaviour. We might therefore expect that information decoding should be more tightly yoked to language performance than activation. Indeed, a recent neurocomputational model of spoken language production and its recovery following virtual lesioning found that the pattern of multivariate information across units in a recovering model was more closely associated with model performance during recovery than univariate activation in those same units; a prediction which has not been empirically tested (Chang & Lambon Ralph, 2020). Nevertheless, only a handful of case reports and studies have utilised MVPA in PSA (Fischer-Baum et al., 2017; Lee et al., 2017; Li et al., 2021).

Having identified which neural regions become more or less functionally involved in language post-stroke, the subsequent key question is: are these regions performing domain-general or language-specific processing (Geranmayeh et al., 2014). 'Domain-general' regions are activated during a wide variety of both language and non-language tasks that are cognitively demanding (Duncan, 2010; Fedorenko et al., 2013) and are thought to mediate executive processes (e.g., attention, strategy selection and performance monitoring) (Fedorenko et al., 2013). Right hemisphere activation during language has historically been assumed to represent the

involvement of novel language-specific quiescent homologues via degeneracy (Blank et al., 2003; Price & Friston, 2002) but it has more recently been appreciated that such regions might instead be subserving domain-general cognitive control processes that are upregulated in PSA following damage to domain-specific language regions (Geranmayeh et al., 2014; Geranmayeh et al., 2017). A recent large-scale meta-analysis of functional neuroimaging literature in PSA found that executive-control related regions of the right anterior insula and inferior frontal gyrus (IFG) were more likely to be activated during language in PSA than controls (Stefaniak et al., 2021). However, while domain-general executive regions are likely involved during difficult language tasks in healthy participants and in PSA (Fedorenko et al., 2013; Stefaniak et al., 2021), it is controversial as to whether domain-general regions are routinely involved during naturalistic language tasks with minimal explicit task demands (Fedorenko et al., 2011; Fedorenko et al., 2012). In the context of information processing, it is also unclear what function such domain-general regions serve during language in PSA and/or intact language.

The present study therefore explored the neural basis of aphasia recovery and of the role performed by domain-general regions, by performing a language MVPA-functional magnetic resonance imaging (fMRI) experiment in a cohort of participants with chronic PSA and controls using a naturalistic language task with minimal explicit demands and thus no expectation that domain-general executive regions should be required. We identified regions in which language information content was positively associated with out-of-scanner language performance, and assessed whether this association between information content and language performance differed between PSA and controls. If regions in which decoding correlated with language had increased decoding accuracy in PSA relative to controls, this would suggest they served a compensatory role to aid aphasia recovery. If such regions had decreased decoding accuracy in PSA relative to controls, this would suggest they contributed to the language deficit through a novel form of diaschisis. We assessed whether univariate language activation extracted from the above regions was associated with out-of-scanner language performance. If not, this would suggest that multivariate information content correlates more closely with recovery than activation-based metrics in at least some brain regions, as predicted by the computational model (Chang & Lambon Ralph, 2020). Finally, we assessed whether these ‘decoding’ regions were language-specific or domain-general by extracting from them univariate activation during a non-language cognitively demanding task, and determining whether this too was associated with language performance post-stroke (Geranmayeh et al., 2014). If such 'decoding' regions were also domain-general executive, this would suggest that domain-general regions are functionally involved during language even during naturalistic language tasks with minimal explicit demands,

and are capable of processing and/or detecting high level language information that correlates with language performance post stroke.

Materials and methods

Participants

PSA participants were recruited from the community throughout the East of England while healthy controls were recruited from the volunteer panel of the MRC Cognition and Brain Sciences Unit. Inclusion criteria for patients were: at least 6 months post left hemispheric stroke; stroke detectable on T_1 MRI scan; and no previous strokes or other neurological conditions. Inclusion criteria for all participants were: right-handed; native English speakers; without contraindication to MRI scanning; with normal or corrected-to-normal hearing and vision; and able to complete an adequate listening task scan. Informed consent was obtained from all participants according to the Declaration of Helsinki under approval from the local NHS research ethics committee. Thirty controls and 26 patients were recruited, completed neuropsychological testing and attended the MRI scan. Of these, two patients had to be excluded from the primary analysis because we were unable to obtain an adequate listening task scan (they did not button press during the listening task to indicate they were attending to the auditory stimuli). This left 24 patients and 30 controls in the final analysis. Of these, two patients and one control were unable to produce usable data for the 'pattern-matching' task (one patient and one control due to a technical issue affecting the timing of stimulus presentation, one patient due to fatigue causing us to stop the task prematurely). Analyses on 'pattern-matching' data therefore included 22 patients and 29 controls.

The PSA and control groups were matched for age (mean 59.8 [SD 14.6] years in PSA vs. 62.5 [SD 5.5] in controls; *t*-test, t_{28}=-0.86, p_{uncorr}=0.40), years of education (mean 14.9 [SD 3.5] years in PSA vs. 16.3 [SD 3.0] in controls; *t*-test, t_{46}=-1.62, p_{uncorr}=0.11), handedness (mean Edinburgh Handedness Inventory score 86.0 [SD 23.9] in PSA vs. 90.0 [SD 12.5] in controls; *t*-test, t_{33}=-0.74, p_{uncorr}=0.47) and sex (17/24 males in stroke survivors vs 22/30 males in controls; Chi-Square test, χ^2_1=0.04, p_{uncorr}=0.84). Demographic and clinical variables for the PSA and control groups are shown in Supplementary Table S4.1.

Neuropsychological tests

All participants completed a data-driven reduced battery consisting of 14 neuropsychological tests that was previously developed in an independent sample of chronic PSA patients and designed to provide the optimal balance between testing time and ability to sensitively detect even mild deficits on underlying neuropsychological components (Halai, De Dios Perez, et al., 2020). See Supplementary Methods for further details.

Statistical analysis

Group differences between patients and controls were assessed using independent sample *t*-tests, not assuming equal variances, for continuous variables and Chi-Square test for categorical variables. We defined statistical significance as two-tailed $p<0.05$. Multiple comparisons were corrected for using the False Discovery Rate (FDR) $q<0.05$ (Benjamini & Hochberg, 1995). We performed varimax-rotated principal component analysis (PCA) on the correlation matrix of the combined neuropsychological data of PSA and controls in order to reduce these 14 test scores to a smaller number of underlying dimensions. Scores from principal components (PCs) with an eigenvalue greater than one were taken to be estimates of underlying cognitive components. We used SPSSv25 for group comparisons and the PCA; all other analyses were performed in Matlab 2018a.

Voxel Based Correlational Methodology

Bias corrected T_1-weighted structural images were segmented, and the resultant grey matter tissue probability maps normalised, using the ALIv3 modified unified segmentation-normalisation procedure (Seghier et al., 2008) implemented in Statistical Parametric Mapping (SPM12, www.fil.ion.ucl.ac.uk/spm/). The grey matter tissue probability maps in MNI space were smoothed using an 8mm FWHM Gaussian kernel and taken to be voxelwise estimates of grey matter tissue integrity. These 'first-level' grey matter tissue integrity maps of both patients and controls were entered into combined mass univariate 'second-level' voxel-based correlational methodology (VBCM) analyses (Tyler et al., 2005)**,** i.e. correlated with principal component language scores while controlling for lesion volume, using the CANlab Robust Regression Toolbox (Wager et al., 2005) (https://github.com/canlab/RobustToolbox) and SPM12. Correction for multiple comparisons used voxelwise FDR $q<0.05$ (two-sided) with a minimum of 20 significant voxels per cluster.

Functional MRI design

All participants completed a scanning session involving as many of three cognitive tasks that they were able to complete. All participants were trained on all tasks outside the scanner before the session. No information was given as to the aim of the study or what we were expecting from each cognitive task. During the session, stimuli were presented using Eprime2 run on Windows 10. All participants held a button press box in their left hand. Audio files were presented though OptoActive II Active Noise Cancelling Headphones. Visual stimuli were projected onto a screen visible through a headcage-mounted mirror. All fMRI tasks used a blocked fMRI design.

Listening task

There were five 'listening' runs with each run lasting 240s (134 TRs). Runs were presented in a counterbalanced order across participants. During each run, participants passively listened to blocks of three audio conditions which were presented in a pseudorandomised order: normal intelligible speech (4 blocks); pseudoword strings (4 blocks); and rotated-vocoded versions of the pseudoword strings (4 blocks). Each audio block lasted 12 seconds and contained 4x3-second trials, during each trial one audio file was played. In between each audio block was a rest block (12 rest blocks per run, total 87 seconds rest per run). Randomly interspersed between blocks were 3x3-second 'button press' trials per run, during which participants heard an instruction to press the button; this ensured participants were alert and attending to the auditory stimuli. Two participants were excluded because they did not button press during the listening

task. A list of sentences are available in Supplementary Table S4.2.

Sentences were all canonical, subject-verb-object and 5-6 words in length. Half of the sentences (10 'high cloze sentence' blocks containing 40 sentences over the task, 2 blocks per run) had high semantic predictability endings (e.g. 'the pilot flies the plane', 'the gardener mows the lawn', 'the cat chases the mouse'). The remaining half of the sentences (10 'low cloze sentence' blocks containing 40 sentences over the task, 2 blocks per run) each ended in the same final word as a corresponding high cloze sentence, but with the preceding context changed such that the final word had a low semantic predictability ('the bike hits the plane', 'the rocket reaches the lawn', 'the nurse dresses the mouse'). After the listening task, controls were asked to rate how predictable the last word of each sentence was on a scale from 1 (low predictability) to 10 (high predictability) to confirm that sentences were appropriately categorised as high or low cloze. The primary analysis for this paper, a contrast between low and high cloze sentence blocks, was expected to identify regions involved in high-level semantic control (Jackson, 2021) processing during naturalistic sentence listening in the absence of explicit, artificial task demands that would be expected to involve domain-general executive regions.

Contrasts between sentences and pseudoword strings or rotated-vocoded noise were designed for use in an additional study and were not used in the present paper. As part of an additional experiment not included in the present paper, controls performed the five runs of the listening task twice; one time with 'clear' auditory stimuli, the other time with 'degraded' stimuli that were noise-vocoded to have reduced intelligibility. Half of the controls performed the 'degraded listening' task before the 'clear listening task'; an independent samples *t*-test confirmed that the 'clear first' controls did not have significantly different activation/information content to the 'degraded' first controls, during the 'clear listening' task, and therefore the 'clear first' and 'degraded first' controls were grouped together in this paper.

Pattern-matching task

A second MRI task involved visuospatial pattern-matching. Each trial started with a 500ms fixation cross before three black-and-white checkerboard patterns were presented simultaneously, one pattern at the top of the screen and two patterns at the bottom. The participant had three seconds to press the left or right button to indicate whether the left or right pattern at the bottom was the same as the pattern at the top. Half of the trials were 'easy' as they used 3x3 checkerboards. The other half of the trials were hard as they used 4x4 checkboards. Example of 'easy' and 'hard' patterns are available in Supplementary Fig. S4.1. There were 4x3.5s trials per 14s block. A single run lasted 542s (303 TRs) and contained 14 'easy' blocks, 14 'hard' blocks

and 15 blocks of rest (150s divided into 15 blocks of 8-12s each). The contrast of 'hard>easy' pattern-matching blocks was expected to identify regions involved in difficult visuospatial tasks. Importantly, if regions involved in high vs low cloze sentence processing (from the 'listening' MRI task) were significantly more activated in hard than easy pattern-matching, it would suggest such regions were involved in domain-general cognitive-difficulty processing as well.

Other tasks

As part of an additional study not included in the present paper, participants performed a third 'repetition' MRI task.

Functional MRI first-level analysis – mass univariate activation

For the 'activation' analysis, preprocessed EPIs were smoothed with a 2mm FWHM Gaussian kernel and then entered into a mass-univariate fixed-effect first level analysis in SPM12 using the general linear model (Friston et al., 1995) in which each block was modelled as an epoch convolved with the canonical haemodynamic response function. For the 'listening' MRI task, separate regressors were modelled for blocks of: high cloze sentences; low cloze sentences; pseudoword strings; rotated-vocoded noise; rest; and button-press trials. Each participant's five listening runs were modelled as separate sessions in the same first-level design matrix. For the 'pattern-matching' MRI task, separate regressors were modelled for blocks of: hard patterns; easy patterns; and rest. The resultant 'activation contrast estimate' maps were entered into second-level analyses (see below).

Functional MRI first-level analysis – multivariate information

For the 'information' analysis, preprocessed EPIs from the 'listening' MRI task were left unsmoothed but otherwise entered into the same mass-univariate first level analysis as the 'activation' analysis above. Normalised but unsmoothed beta images for 'high cloze sentences' and 'low cloze sentences' were submitted to 'The Decoding Toolbox' v3.997 (Hebart et al., 2014). We performed a whole-brain roaming searchlight (Kriegeskorte et al., 2006) first-level MVPA using a 'pattern-correlation classifier' (Haxby et al., 2001; Misaki et al., 2010) with an 8mm radius spherical searchlight and a leave-one-run-out cross validation design. We used a 'pattern-correlation classifier' (Haxby et al., 2001; Misaki et al., 2010) so that information content decoding in the present study would be independent of differences in mean univariate activation between high and low cloze sentences, and thus could be ascribed solely to differences in the multivariate pattern of distributed activation between conditions. Briefly, the multivoxel pattern of betas for a category (e.g., 'high cloze sentences') from a left-out 'test' run was correlated

against the pattern of betas for (i) the same category, and (ii) the different category (e.g., 'low cloze sentences'), from the other four 'training' runs. The category with the higher Pearson correlation coefficient (same vs. different category) was the classification. This was repeated with each of five runs left-out in turn, and the 'classification accuracy-minus-chance' level ascribed to the voxel at the centre of the searchlight. This was repeated in all voxels throughout the brain mask. The resultant 'classification accuracy-minus-chance' maps were smoothed with a 2mm FWHM Gaussian kernel and entered into second-level analyses (see below).

Functional MRI second-level analysis

First-level parameter estimate maps were entered into group-level random-effects mass univariate 'second-level' analyses. Second-level analyses were performed for two 'activation' contrasts ('low>high cloze sentences' from the 'listening' task, referred to as 'sentence cloze activation', and 'hard>easy patterns' from the 'pattern-matching' task, referred to as 'pattern-matching activation') and one MVPA 'information' analysis ('high vs. low cloze sentences' from the 'listening' task, referred to as 'sentence cloze decoding'). Second-level analyses used the CANlab Robust Regression Toolbox (Mumford, 2017; Wager et al., 2005; Woo et al., 2014) (https://github.com/canlab/RobustToolbox) and SPM12. Correction for multiple comparisons used voxelwise FDR $q<0.05$ (two-sided) with a minimum of 20 significant voxels per cluster.

To identify regions in which activation or information content differed significantly at the group level between patients and controls, 'univariate parameter estimate' and 'multivariate parameter estimate' maps were entered into second-level independent sample *t*-tests comparing patients vs controls. Since we did not identify any regions of significantly different univariate or multivariate parameter estimates between patients and controls, we combined all participants together for subsequent second-level analyses.

To identify regions of significant sentence cloze activation, pattern-matching activation or sentence cloze decoding, second-level one sample *t*-tests identified voxels in which parameter estimate values were significantly different from zero. To investigate regions of significant sentence cloze decoding further, we used custom Matlab script to extract mean first-level analysis parameter estimate values from each of these clusters and compared extracted mean values against zero using one-sample *t*-tests.

To identify regions in which sentence cloze decoding was associated with out-of-scanner language performance, sentence cloze decoding 'multivariate parameter estimate' maps were entered into second-level analyses with PC1 score included as a regressor of interest. To explore further the reason why sentence cloze decoding might be associated with language performance,

we used custom Matlab script to extract mean first-level analysis parameter estimate values from each of these clusters, both as individual clusters ROIs and for all clusters together as a single ROI. We compared extracted mean sentence cloze decoding, sentence cloze activation and pattern-matching activation: against zero using one sample *t*-tests; between patients and control groups using ANCOVAs controlling for grey matter tissue integrity; for associations with PC1 score using robust regression controlling for grey matter tissue integrity; and for a significant interaction between participant group*parameter estimate in the association with PC1 score using robust regression controlling for grey matter tissue integrity.

Power analysis

No study had performed an MVPA experiment in a cohort of patients with PSA before the start of this study and so it was not possible to perform a power analysis for language information content. A power analysis using NeuroPowerTools.org (http://neuropowertools.org/neuropower/neuropowerstart/) on previously collected fMRI data from a similar passive sentence listening experiment in healthy controls estimated a sample size of 22 participants would have power of 0.79 at detecting univariate language activation using FDR whole-brain multiple comparisons correction (Halai et al., 2015).

Data availability

The data that support the findings of this study are potentially available upon request to the corresponding author.
See the Supplementary material for additional methodological details.

Results

Neuropsychological data

Neuropsychological test scores for the PSA and control groups are shown in Supplementary Table S4.3. At the group level, patients with PSA were significantly worse than controls on: immediate word repetition; immediate non-word repetition; the CSB picture naming test; the BNT; forward and backward digit spans; the Cambridge synonym judgement test; the CAT spoken sentence comprehension test; the RCPM; the BSAT; and the number of tokens, WPM and MLU from the 'Cookie Theft' description (Supplementary Table S4.4). Only CSB word-to-picture matching was not significantly worse in PSA than controls (Supplementary Table S4.4).

Principal Component Analysis of the neuropsychological data

We performed varimax-rotated PCA on the correlation matrix of neuropsychological test scores of the PSA (n=24) and control (n=30) participants combined (n=54). The Kaiser-Meyer-Olkin value was 0.82, indicating adequate sampling (Kaiser, 1974). Bartlett's test of sphericity was significant (approximate χ^2_{91}=681, p=4.8x10^{-91}), suggesting these data could be factorised. Two rotated PCs with eigenvalues greater than one explaining 39.7% (PC1) and 28.4% (PC2) of the variance were obtained. PC1 was loaded onto primarily by: immediate word and non-word repetition; the CSB picture naming test and BNT; CSB word-to-picture matching; forward digit span; and CAT spoken sentence comprehension (Table 4.1). PC1 was interpreted as representing phonological-semantic ability. PC2 was loaded onto primarily by tests of executive function (RCPM, BSAT) and speech quanta measures extracted from the BDAE 'Cookie Theft' picture description (Table 1). PC2 was interpreted as representing executive-fluency ability. The PSA group had significantly lower PC1 (mean -0.48 [SD 1.35] in PSA vs 0.39 [SD 0.20] in controls; t-test, t_{24}=-3.1, p_{uncorr}=0.005) and PC2 scores (mean -0.67 [SD 1.00] in PSA vs 0.54 [SD 0.60] in controls; t-test, t_{36}=-5.2, p_{uncorr}=8x10^{-6}) than the control group. However, their distributions showed graded overlap between patients and controls (Supplementary Fig. S4.2, Supplementary Table S4.5), in keeping with patients being a heterogenous group showing marked inter-individual variability in language performance ranging from near-normal to significantly impaired.

Neuropsychological test	Component loadings	
	PC1	PC2
Immediate Word Repetition	**0.93**	0.10
Immediate Non-word Repetition	**0.81**	0.40
CSB Picture Naming Test	**0.92**	0.27
Boston Naming Test	**0.83**	0.35
CSB Word-Picture Matching	**0.90**	0.01
Forward Digit Span	**0.67**	0.57
Backward Digit Span	0.48	0.58
Cambridge Synonym Judgement Test	0.31	0.58
CAT Spoken Sentence Comprehension Test	**0.83**	0.39
Raven's Coloured Progressive Matrices	0.18	**0.75**
Brixton Spatial Anticipation Test	0.17	**0.65**
'Cookie Theft' Number of tokens	0.07	**0.68**
'Cookie Theft' Words Per Minute	0.40	**0.65**
'Cookie Theft' Mean Length of Utterance	0.10	**0.78**

Table 4.1: Component matrix of neuropsychological scores from PCA performed in PSA and controls combined. Varimax rotated principal component analysis was performed on the combined neuropsychological data of patients with post-stroke aphasia and controls. The loading of each score onto each rotated principal component is shown. Variables with major loadings (defined as >0.60) are in bold. Immediate word repetition was from the PALPA9; immediate non-word repetition was from the PALPA8. Abbreviations: CAT = Comprehensive Aphasia Test; CSB = Cambridge Semantic Battery; PALPA = Psycholinguistic Assessments of Language Processing in Aphasia; PCA = Principal Component Analysis; PC = Principal Component.

Voxel Based Correlational Methodology

The lesion overlap mask of the PSA group involves most of the left hemisphere (Supplementary Fig. S4.3). Scores from principal components (PCs) were regressed against voxelwise grey matter tissue integrity, controlling for lesion volume, through VBCM in patients and controls combined.

Lower PC1 score was associated with lower grey matter tissue integrity in expected canonical left fronto-temporo-parietal language regions, with the largest cluster's peak in the left middle temporal gyrus (MTG) (Fig. 4.1A, Supplementary Table S4.6). As expected, the PSA group had significantly lower mean grey matter tissue integrity in the PC1 cluster than the control group (mean 0.44 [SD 0.12] in PSA vs 0.54 [SD 0.04] in controls; *t*-test, t_{27}=-4.2, p_{uncorr}=2x10^{-4}). However, their distributions showed graded overlap between patients and controls (Fig. 4.1B), in keeping with patients being a heterogeneous group showing marked inter-individual variability in the degree of damage to this region. This overlap suggested against treating patients as a homogeneous group that is completely separate to controls. PC2 did not have a neural correlate on VBCM. We therefore focused our subsequent analyses on PC1.

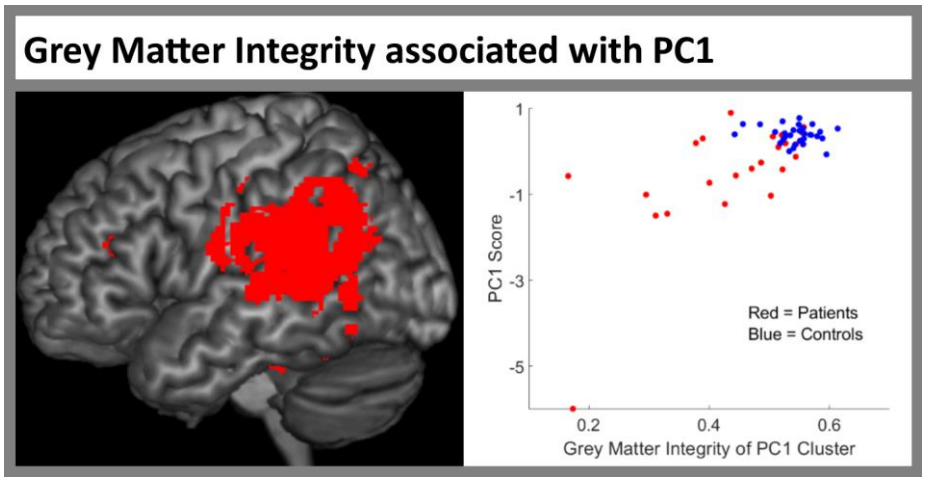


Figure 4.1: Grey matter integrity associated with language. Left: clusters (red) in which lower grey matter tissue integrity was associated with worse language performance, as measured by PC1 score, obtained from Voxel Based Correlational Methodology in patients and controls combined (n=54) controlling for lesion volume. Statistical thresholding used voxelwise FDR q<0.05 (two-sided) with a minimum of 20 significant voxels per cluster. Right: scatter plot showing the association between mean grey matter tissue integrity extracted from the left panel's clusters (x-axis), and PC1 score (y-axis), in patients (red) and controls (blue). Abbreviations: FDR = 'False Discovery Rate'; PC = 'Principal Component'.

Functional MRI

26 patients and 30 controls attended the MRI scan. Of these, 24 patients and 30 controls had an adequate 'listening' scan, and 22 patients and 29 controls had an adequate 'pattern-matching' scan. This meant that a majority of patients were able to perform our low-demand passive sentence listening task (24/26 patients scanned) and our non-verbal 'pattern-matching' task (22/26 patients scanned). This suggests that it is feasible to perform task-based fMRI in patients with PSA.

Participants who completed the 'listening task' had a mean button press response rate of 14.1 (SD 2.1) out of 15, suggesting the auditory stimuli were well attended to. Patients and controls did not have significantly different button press response rates (mean 13.8 [SD 2.9] out of 15 in patients vs 14.2 [SD 1.2] out of 15 in controls; *t*-test, t_{26}=-0.66, p_{uncorr}=0.52). After the scan, controls rated 'high cloze sentences' as having significantly more predictable endings than 'low cloze sentences' (mean 9.2 (SD 0.4) in high cloze sentences vs 2.1 (SD 0.7) in low cloze sentences; independent *t*-test, t_{64}=57.3, p_{uncorr}=1.4x10^{-56}), suggesting this manipulation should enable us to identify regions involved in high-level semantic processing during naturalistic sentence comprehension.

For the 51 participants who completed the pattern-matching tasks, mean response rate during the pattern-matching task was 94.5% (SD 7.8%). Compared to 'easy' pattern-matching trials, 'hard' trials had significantly lower accuracy (mean 96.3% [SD 7.8%] correct in 'easy' vs 80.0% [SD 17.4%] correct in 'hard'; paired *t*-test, t_{50}=8.67, p_{uncorr}=2x10^{-11}) and significantly longer response times (mean 1234ms [SD 274ms] in 'easy' vs 2078ms [250ms] in 'hard'; paired *t*-test, t_{50}=-28.65, p_{uncorr}=1x10^{-32}), suggesting that the 'hard>easy' pattern-matching contrast extracted task difficulty.

We did not identify any regions of significantly different sentence cloze activation, pattern-matching activation or sentence cloze decoding between patients and controls on independent sample *t*-tests. Additionally, we did not identify any regions of significantly different sentence cloze activation or sentence cloze decoding between controls who performed the clear listening task first and controls who performed the vocoded listening task first, on independent sample *t*-tests. We therefore combined all participants (patients, clear-first controls and vocoded-first controls) into the same group for subsequent second-level analyses.

Functional MRI – sentence cloze activation

One sample *t*-test

Seven clusters of significant 'sentence cloze activation' were present in core regions of the 'semantic control network' (Jackson, 2021) (including bilateral IFG opercularis/triangularis/orbitalis and cerebellum, left MTG and medial superior frontal gyrus (SFG)) (Supplementary Table S4.7, Fig. 4.2A). This was expected given that comparing high against low cloze sentences should utilise semantic control as it requires the ability to manipulate semantic knowledge of the terminal word within the preceding sentence's context. 12 clusters of significant 'sentence cloze deactivation' were present in core regions of the Default Mode Network (DMN) (including midline precuneus/posterior cingulate cortex (PCC)/anterior

cingulate cortex (ACC)/medial orbital SFG) as well as bilateral superior temporal gyrus (STG)/Heschl's gyrus/rolandic operculum and right MTG/inferior temporal gyrus (ITG) (Supplementary Table S4.7, Fig. 4.2A).

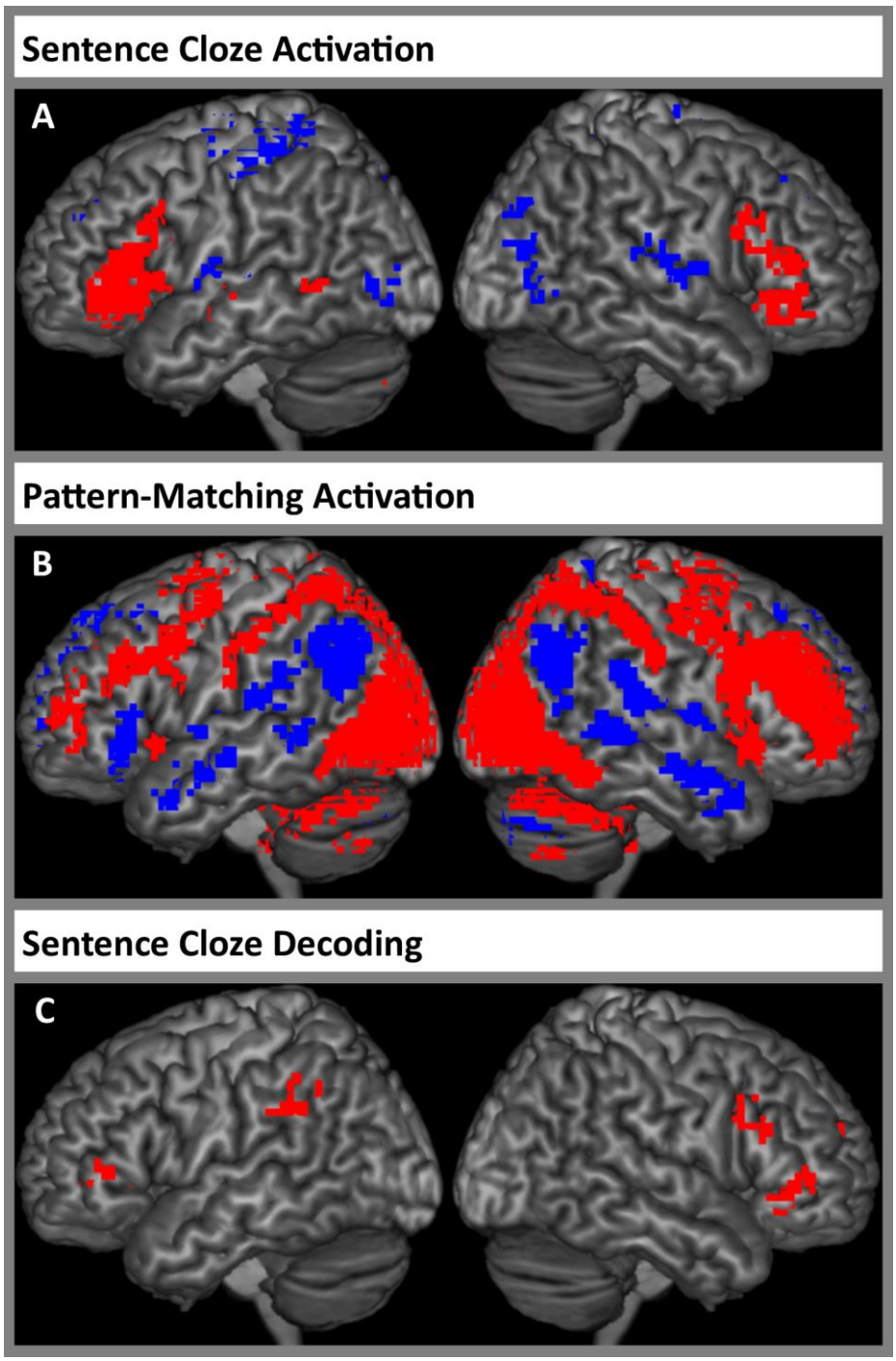


Figure 4.2: Regional activation and information content decoding. (A) Clusters of significant 'low>high' sentence cloze activation (red) or deactivation (blue). **(B)** Clusters of significant 'hard>easy' pattern matching activation (red) or deactivation (blue). **(C)** Clusters of significantly above chance 'high versus low' sentence cloze decoding (red). The left and right columns show the left and right hemispheres, respectively. Robust regression used a second-level one-sample *t*-test on the first-level univariate or multivariate parameter estimate maps of patients and controls combined (n=54). Statistical thresholding used voxelwise FDR q<0.05 (two-sided) with a minimum of 20 significant voxels per cluster. Abbreviations: FDR = 'False Discovery Rate'.

Functional MRI – pattern-matching activation

One sample *t*-test

Two clusters of significant 'pattern-matching activation' were present in regions of the frontoparietal multiple demand network (including bilateral IFG opercularis/triangularis/orbitalis/MFG/dorsolateral SFG/medial SFG/supplementary motor area (SMA)/ACC/median cingulate cortex (MCC)/superior parietal gyrus/insula/rolandic operculum/thalamus/caudate/putamen/cerebellum) (Supplementary Table S4.8, Fig. 4.2B). This was expected as such 'multiple demand' regions are activated across a range of cognitively demanding tasks (Duncan, 2010). 13 clusters of significant 'pattern-matching deactivation' were present in bilateral regions of the DMN (including midline precuneus/medial orbital SFG/ACC/MCC/PCC and bilateral angular gyrus/supramarginal gyrus (SMG)/STG/MTG/temporal pole (Supplementary Table S4.8, Fig. 4.2B).

Functional MRI – sentence cloze information

One sample *t*-test

Nine clusters of significantly above-chance sentence cloze decoding were present in core regions of the 'semantic control network' (Jackson, 2021) (including bilateral IFG triangularis/MFG/dorsolateral SFG, midline ACC/SMA/medial SFG and right IFG opercularis/orbitalis) as well as regions not typically associated with semantic control (Jackson, 2021) such as the left SMG and midline precuneus (Supplementary Table S4.9, Fig. 4.2C). Seven of these nine clusters had significant sentence cloze activation (clusters 1, 5, 7, 9 including the left IFG triangularis, midline SMA/medial SFG and right IFG opercularis/triangularis/orbitalis/MFG) or deactivation (clusters 2, 6, 8 including the midline ACC/precuneus/medial SFG), but two of these nine clusters did not have significant sentence cloze activation/deactivation (clusters 3, 4 including the left SMG and right MFG/dorsolateral SFG/medial SFG) (Supplementary Table S4.9). Thus, significant language information decoding occurred in certain regions from both hemispheres that were not significantly activated or deactivated by the corresponding univariate language contrast and might therefore have been thought of as uninvolved in high-level language processing were activation considered in isolation. This demonstrates the added value of looking at information content decoding when looking at the neural substrates of language in health and PSA.

Three of these nine clusters, all located within the left hemisphere, did not have significant pattern-matching activation/deactivation (clusters 3, 5, 8 including the left IFG triangularis/MFG/dorsolateral SFG/SMG) but six of these nine clusters, all located in midline or right hemisphere regions, had significant pattern-matching activation (clusters 1, 4, 6, 7, 9 including the midline SMA/medial SFG and right IFG

opercularis/triangularis/orbitalis/MFG/dorsolateral SFG/medial SFG) or deactivation (cluster 2 including the medial SFG/ACC) (Supplementary Table S4.9). Thus, midline and right hemisphere regions that performed high-level language-specific computations during naturalistic sentence listening were also domain-general and differentially activated by more or less demanding non-language tasks, despite the 'high vs low cloze sentence' comparison used for information content decoding not requiring an explicit task to be performed and tapping into high-level linguistic comprehension with no obvious requirement for non-language executive function.

Association with out-of-scanner language performance

There were no clusters in which sentence cloze decoding was negatively associated with PC1 score. However, 21 clusters were identified in which sentence cloze decoding was positively associated with PC1 score (Supplementary Table S4.10, Fig. 4.3A). These clusters were bilateral albeit left lateralised and included extensive regions of the left hemisphere language network (Fedorenko et al., 2011) (including left IFG opercularis/triangularis/orbitalis/MFG/dorsolateral SFG/medial SFG/ACC, temporal pole/MTG/ITG/fusiform gyrus, SMG/inferior parietal gyrus, insula, putamen and cerebellum), as well as midline (precuneus/medial SFG), and, to a lesser extent, right fronto-temporo-parietal cortex (including right IFG opercularis/triangularis/orbitalis/dorsolateral SFG/medial SFG/MTG/STG/rolandic operculum/insula/postcentral gyrus/precuneus/lingual gyrus) (Supplementary Table S4.10, Fig. 4.3A).

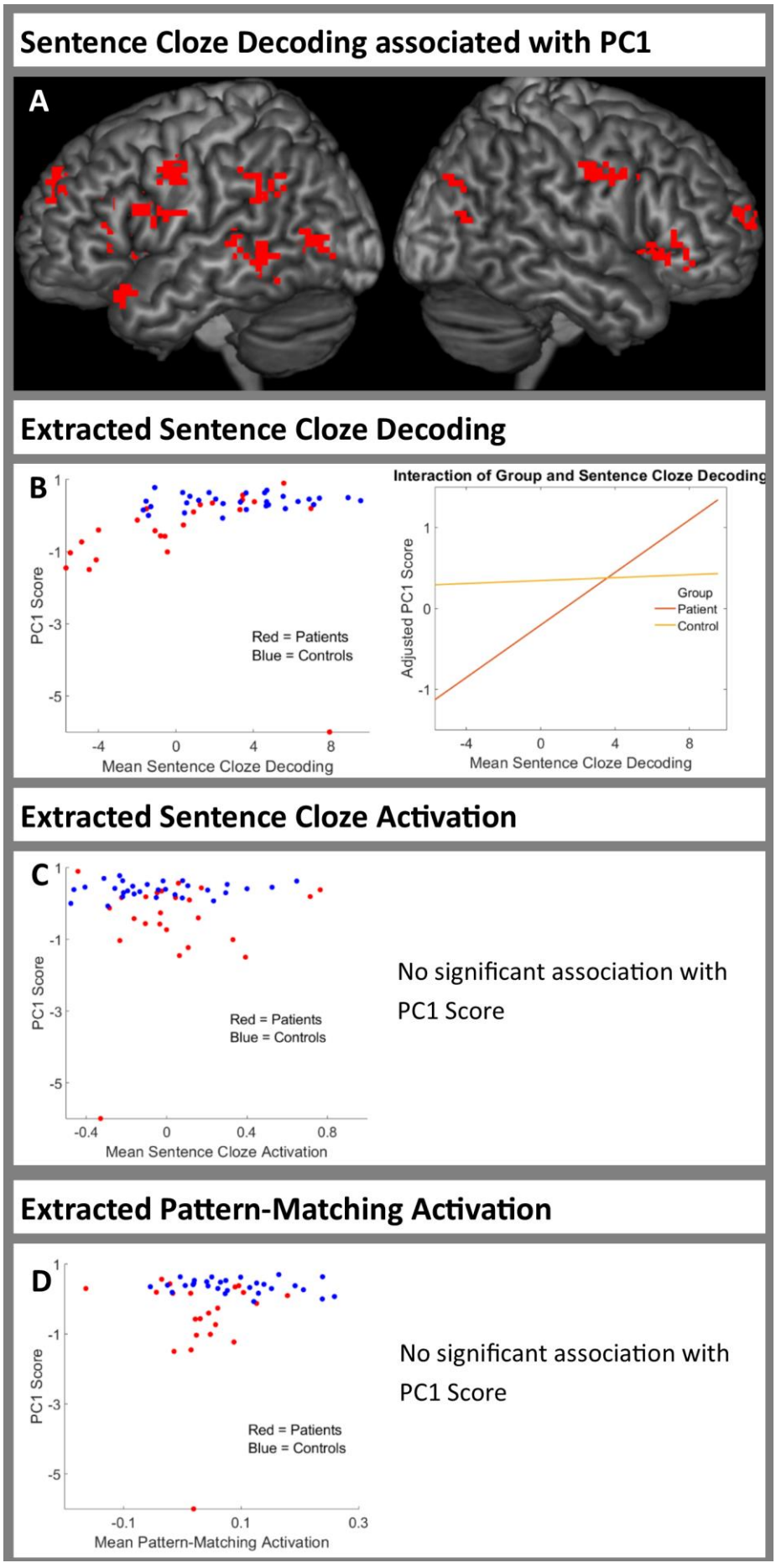


Figure 4.3: Regions in which sentence cloze decoding was positively associated with language. (A) Clusters in which 'high versus low' sentence cloze decoding was positively associated with PC1 score using a second-level robust regression analysis on the first-level multivariate parameter estimate maps of patients and controls combined (n=54). The left and right columns show the left and right hemispheres, respectively. Statistical

thresholding used voxelwise FDR q<0.05 (two-sided) with a minimum of 20 significant voxels per cluster. **(B)** Mean 'high vs low' sentence cloze decoding values were extracted from all clusters in (A) for each participant and entered into robust regression analyses against PC1, controlling for mean grey matter tissue integrity and including a group*'sentence cloze decoding' interaction term. Left: scatter plot showing the association between mean extracted sentence cloze decoding (x-axis), and PC1 score (y-axis), in patients (red) and controls (blue). Right: plot showing the model-predicted significant group*'sentence cloze decoding' interaction such that lower sentence cloze decoding was associated with worse language performance in patients (red) but not in controls (yellow). **(C)** Mean 'low>high' sentence cloze activation values were extracted from all clusters in (A) for each participant and entered into robust regression analyses against PC1, controlling for mean grey matter tissue integrity and including a group*'sentence cloze activation' interaction term. Left: scatter plot showing the association between mean extracted sentence cloze activation (x-axis), and PC1 score (y-axis), in patients (red) and controls (blue). Right: there was no significant association between mean sentence cloze activation and PC1 score, nor was the group*'sentence cloze activation' interaction significant. **(D)** Mean 'hard>easy' pattern-matching activation values were extracted from all clusters in (A) for each participant and entered into robust regression analyses against PC1, controlling for mean grey matter tissue integrity and response rate and including a group*'pattern-matching activation' interaction term. Left: scatter plot showing the association between mean extracted pattern-matching activation (x-axis), and PC1 score (y-axis), in patients (red) and controls (blue). Right: there was no significant association between mean pattern-matching activation and PC1 score, nor was the group*'pattern-matching activation' interaction significant. Abbreviations: FDR = 'False Discovery Rate'; PC = 'Principal Component'.

Mean sentence cloze decoding and grey matter tissue integrity were extracted from each of these 21 clusters (in which sentence cloze decoding was positively associated with PC1 score), individually as well as from all 21 clusters together. Controlling for grey matter tissue integrity, there was a significant group*'sentence cloze decoding' interaction in 17 of these 21 clusters such that lower sentence cloze decoding was associated with worse language performance in patients but not in controls (clusters 1-10, 12, 14-15, 17-19, 21 including the bilateral IFG opercularis/triangularis/orbitalis/dorsolateral SFG/medial SFG/precentral gyrus/MTG/insula, left MFG/SMG/inferior parietal gyrus/temporal pole/ITG/fusiform gyrus/putamen/cerebellum and right postcentral gyrus/precuneus/lingual gyrus) (Supplementary Table S4.10). For display purposes, we confirmed that this interaction remained significant (p=0.002) when all clusters were considered as a single ROI (Table 4.2, Fig. 4.3B). Critically, mean sentence cloze decoding was numerically lower in patients than controls in all 21 clusters; significantly lower in four of these 21 clusters (clusters 2, 4, 6, 16 including the left MFG/dorsolateral SFG/medial SFG/ACC and right IFG opercularis/triangularis/orbitalis/MTG/insula); and significantly lower when all clusters were considered as a single ROI ($F(1,51)=5.67$, p=0.02) (Supplementary Table S4.10). Together, these results demonstrate that performance in this post-stroke aphasic group is associated with lower language information content in a distributed set of bilateral regions, even controlling for tissue integrity, and the degree by which language information content decreases is associated with language performance. This provides evidence for the existence of a novel form of diaschisis ('information diaschisis') that might contribute to the language deficit in PSA. These results also implicate language information processing in a distributed set of undamaged right

fronto-temporo-parietal neural regions as being important for maintenance of language performance post-stroke.

Model/variable	B	SE	p-value	Adjusted R^2	N
Analyses with mean 'sentence cloze decoding' extracted from all clusters (Fig. 4.3B)					
Model 1:			$1.5x10^{-5}$*	0.33	54
PC1Score ~ 1 + Mean Sentence Cloze Decoding + ClusterSwc1T1					
Constant	-0.78	0.63	0.22		
Mean Sentence Cloze Decoding	**0.09**	**0.02**	**$6.4x10^{-5}$***		
Mean Grey Matter Tissue Integrity of Clusters	1.63	1.38	0.24		
Model 2:			$2.0x10^{-6}$*	0.43	54
PC1Score ~ 1 + Group + Mean Sentence Cloze Decoding + Group*Mean Sentence Cloze Decoding + ClusterSwc1T1					
Constant	-0.51	0.70	0.47		
Group (Patient or Control)	0.55	0.20	0.009*		
Mean Sentence Cloze Decoding	0.16	0.03	$3.2x10^{-6}$*		
Group*Mean Sentence Cloze Decoding	**-0.15**	**0.05**	**0.002***		
Mean Grey Matter Tissue Integrity of Clusters	0.67	1.59	0.68		

Table 4.2: Clusters in which sentence cloze decoding was positively associated with language score – analyses with extracted 'sentence cloze decoding'. For all clusters in which sentence cloze decoding information content was positively associated with PC1 score (using a regression of PC1 score on the 'multivariate parameter estimate' maps of patients and controls combined (n=54) (Fig. 4.3)), we extracted the mean sentence cloze decoding and mean grey matter tissue integrity. We entered these extracted mean multivariate parameter estimates into robust regression analyses predicting PC1 score, controlling for mean grey matter tissue integrity of the cluster ROIs, with (Model 2) or without (Model 1) a group*mean parameter estimate interaction term. 'ClusterSwc1T1' is the mean grey matter tissue integrity of all clusters combined. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; SE=Standard Error of regression coefficient.

To confirm whether multivariate information content provides additional, complementary data to univariate activation at explaining language performance post-stroke, mean sentence cloze activation was extracted from each of these 21 clusters (in which sentence cloze decoding was positively associated with PC1 score), individually as well as from all 21 clusters together. Controlling for mean grey matter tissue integrity, mean sentence cloze activation was not significantly different in patients than controls in any of these 21 clusters (all $pFDR \geq 0.40$) nor when all clusters were considered as a single ROI ($F(1,51)=0.88$, $p=0.35$); mean sentence cloze activation was not significantly associated with PC1 score in any of these 21 clusters individually

(all *pFDR*=0.99); and there was a significant group*'sentence cloze activation' interaction in only 3 of these 21 clusters (clusters 4-5, 11 in the midline precuneus and right precentral gyrus/postcentral gyrus/MTG) (Supplementary Table S4.11). For display purposes, we confirmed this group*'sentence cloze activation' interaction remained non-significant (*p*=0.76) when all clusters were considered as a single ROI (Table 4.3, Fig. 4.3C). Thus, in a distributed set of bilateral regions, diaschisis of multivariate language information content contributed to the language deficit in PSA, and multivariate language information content was important for maintenance of language performance post-stroke, independently of univariate activation.

Model/variable	B	SE	p-value	Adjusted R^2	N
Analyses with mean 'sentence cloze activation' extracted from all clusters (Fig. 4.3C)					
Model 1:			0.0005*	0.23	54
PC1Score ~ 1 + Mean Sentence Cloze Activation + ClusterSwc1T1					
Constant	-1.03	0.58	0.08		
Mean Sentence Cloze Activation	**0.07**	**0.24**	**0.76**		
Mean Grey Matter Tissue Integrity of Clusters	2.78	1.25	0.03*		
Model 2:			0.009*	0.18	54
PC1Score ~ 1 + Group + Mean Sentence Cloze Activation + Group*Mean Sentence Cloze Activation + ClusterSwc1T1					
Constant	-0.95	0.73	0.20		
Group (Patient or Control)	0.47	0.18	0.01*		
Mean Sentence Cloze Activation	0.30	0.44	0.50		
Group*Mean Sentence Cloze Activation	**-0.18**	**0.60**	**0.76**		
Mean Grey Matter Tissue Integrity of Clusters	1.82	1.66	0.28		

Table 4.3: Clusters in which sentence cloze decoding was positively associated with language score – analyses with extracted 'sentence cloze activation'. For all clusters in which sentence cloze decoding information content was positively associated with PC1 score (using a regression of PC1 score on the 'multivariate parameter estimate' maps of patients and controls combined (n=54) (Fig. 4.3)), we extracted the mean sentence cloze activation and mean grey matter tissue integrity. We entered these extracted mean univariate parameter estimates into robust regression analyses predicting PC1 score, controlling for mean grey matter tissue integrity of the cluster ROIs, with (Model 2) or without (Model 1) a group*mean parameter estimate interaction term. 'ClusterSwc1T1' is the mean grey matter tissue integrity of all clusters combined. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; SE=Standard Error of regression coefficient.

To investigate whether these 21 clusters (in which sentence cloze decoding was positively associated with PC1 score) were language-specific or domain-general, mean pattern-matching activation was extracted from each of them individually as well as from all 21 clusters together. There was significant pattern-matching activation in 12 of these 21 clusters in controls (clusters 1, 4-6, 8-12, 17, 20, 21 including the left IFG opercularis/triangularis/orbitalis/MFG/dorsolateral SFG/precentral gyrus/insula/MTG/fusiform gyrus/putamen/cerebellum, right IFG opercularis/triangularis/orbitalis/precentral gyrus/postcentral gyrus/insula/rolandic operculum/MTG) (Supplementary Table S4.12). Thus, most regions in which language

information content correlated with language performance were domain-general and responsive to increased cognitive-demand. However, mean pattern-matching activation was not significantly associated with PC1 score in any of these 21 clusters (all $pFDR \geq 0.95$); and there was not a significant group*'pattern-matching activation' interaction in any of these 21 clusters (all $pFDR \geq 0.14$) (Supplementary Table S4.12, Fig. 4.3D). For display purposes, we confirmed that this interaction remained non-significant ($p=0.54$) when all clusters were considered as a single ROI (Table 4.4, Fig. 4.3D). Thus, most regions in which language information content correlated with language performance were domain-general; however, the activation of these regions to non-language cognitively demanding tasks or to low-demand language tasks was not associated with language performance. Rather, it was specifically the amount of language information content within these distributed, bilateral, domain-general regions that was associated with language performance in PSA.

Model/variable	B	SE	p-value	Adjusted R^2	N
Analyses with mean 'pattern-matching activation' extracted from all clusters (Fig. 4.3D)					
Model 1:			$1.8x10^{-5}$*	0.37	51
PC1Score ~ 1 + Mean Pattern-Matching Activation + ClusterSwc1T1 + Response Rate					
Constant	-5.13	1.01	$6.6x10^{-6}$*		
Mean Pattern-Matching Activation	**-0.15**	**0.92**	**0.87**		
Mean Grey Matter Tissue Integrity of Clusters	2.74	1.46	0.07		
Response Rate	0.04	0.01	$7.6x10^{-5}$*		
Model 2:			0.0005*	0.31	51
PC1Score ~ 1 + Group + Mean Pattern-Matching Activation + Group* Mean Pattern-Matching Activation + ClusterSwc1T1 + Response Rate					
Constant	-4.45	1.32	0.002*		
Group (Patient or Control)	0.27	0.25	0.29		
Mean Pattern-Matching Activation	-1.75	1.91	0.36		
Group*Mean Pattern-Matching Activation	**1.45**	**2.35**	**0.54**		
Mean Grey Matter Tissue Integrity of Clusters	2.08	1.76	0.24		
Response Rate	0.04	0.01	0.005*		

Table 4.4: Clusters in which sentence cloze decoding was positively associated with language score – analyses with extracted 'pattern-matching activation'. For all clusters in which sentence cloze decoding information content was positively associated with PC1 score (using a regression of PC1 score on the 'multivariate parameter estimate' maps of patients and controls combined (n=54) (Fig. 4.3)), we extracted the mean pattern-matching activation and mean grey matter tissue integrity. We entered these extracted mean univariate parameter estimates into robust regression analyses predicting PC1 score, controlling for mean grey matter tissue integrity of the cluster ROIs and for task response rate, with (Model 2) or without (Model 1) a group*mean parameter estimate interaction term. 'ClusterSwc1T1' is the mean grey matter tissue integrity of all clusters combined. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; SE=Standard Error of regression coefficient.

Discussion

This study investigated the neural basis of aphasia recovery by performing one of the first language MVPA experiments in post-stroke aphasia (PSA) to date. The results show: that MVPA can identify more neural regions that are functionally involved in language and its recovery than traditional 'activation-based' approaches; that decoding accuracy becomes associated with language performance post-stroke in bilateral fronto-temporo-parietal cortex;

that a novel form of 'information diaschisis' contributes to the language deficit in PSA; that most of the regions identified by language decoding were domain-general; and that it is specifically the amount of language information content (rather than language or domain-general activation) within these domain-general regions that is associated with language performance in PSA but not controls. We consider these results in a little more detail below.

First, considering the task-based fMRI, we show that it is feasible to perform an MVPA experiment in a cohort of patients with PSA. Following previous inspiration (Leff et al., 2002; Sharp et al., 2010), we specifically designed our language task to be passive and simple so that we could include stroke survivors with as wide a range of severities as possible. More than 90% of stroke patients scanned were able to complete the listening task, suggesting that it could be trialled in a clinical setting.

Second, our results show that there are localised brain regions, in both left and right fronto-temporo-parietal cortex, in which language information is manifest in the distributed pattern of activation across voxels. This demonstrates that language processing is not restricted to 'canonical' language areas in the left hemisphere. Many clusters were part of the 'semantic control network' (Jackson, 2021) (including the bilateral IFG and midline SFG/ACC/SMA), which was expected given that comparing high against low cloze sentences should utilise the ability to manipulate semantic knowledge of the terminal word within the preceding sentence's context (Hoffman & Tamm, 2020; Humphreys & Gennari, 2014; Kielar et al., 2015). However, significant decoding additionally included regions (e.g., left SMG) that were not significantly activated or deactivated by the corresponding univariate language contrast, and consequently might not have been to be involved in high-level language processing. Furthermore, significant decoding was even found in regions of sentence cloze deactivation (e.g., the precuneus), which might otherwise have been thought of as 'actively disinterested' in high-level language processing using 'activation-based' imaging. Since language information content was inferred using a 'pattern correlation classifier', decoding accuracy was independent of differences in mean activation between high and low cloze sentences and thus was not influenced by deactivation per se. This raises the fascinating possibility that both activated and deactivated regions might compute information subserving language, and demonstrates the added value of MVPA when identifying the neural substrates of language in health and PSA.

Third, as well as identifying novel regions functionally involved in language, MVPA identified regions in which decoding accuracy was positively associated with language performance in PSA but not in controls. As above, these regions extended beyond canonical left hemisphere language regions into midline and right fronto-temporo-parietal cortex, and included the right

insula and IFG, regions which were found by a recent meta-analysis to be more likely to be activated during language in PSA than controls (Stefaniak et al., 2021). Nevertheless, univariate language activation in these same regions was not associated with language performance. This shows that MVPA decoding might be useful as a novel biomarker in PSA research that can be more sensitive than equivalent univariate activation contrasts at detecting neural correlates of language performance. It also provides direct experimental confirmation of a prediction from a recent computational model, which hypothesised that information content might be more important than activation during aphasia recovery (Chang & Lambon Ralph, 2020). This suggests that simply activating a neural region is insufficient to contribute to task performance; the region needs to be 'tuned to' and process task-relevant information as well.

Intriguingly, such regions (in which decoding accuracy was positively associated with language performance) did not have increased decoding accuracy in PSA than controls, as would be expected if such regions served an adaptive, compensatory role to aid aphasia recovery. Instead, they had decreased decoding accuracy in PSA than controls. This suggests that stroke-induced local damage can be associated with lower language information content in distant undamaged nodes of the residual language network, and that the degree by which such distributed language information processing is reduced correlates with, and indeed might contribute to, the degree of language impairment. These results are consistent with a novel form of 'diaschisis' (Carrera & Tononi, 2014), which we herein call 'information diaschisis'. This implicates maintenance of language information processing in a distributed set of bilateral fronto-temporo-parietal neural regions as being important for maintenance of language performance post-stroke.

There is ongoing debate as to whether regions of upregulated language activation post-stroke are in fact domain general and engaged because of increased requirements for top-down executive control during language (Geranmayeh et al., 2014; Geranmayeh et al., 2017; Stefaniak et al., 2020). Thus, a final aim of this study was to assess the role of domain-general regions during language in health and PSA recovery. Our sentence cloze decoding comparison did not require an explicit task and did not require participants to behave or pay attention differently to high or low cloze sentences. Indeed, it was not explicitly revealed to participants that a contrast between these sentences was the purpose of the experiment. This task passively tapped into high-level aspects of semantic comprehension with no explicit requirement for non-language executive function (i.e., there was no need for overt decisions, item comparisons, etc.). Nevertheless, midline and right hemisphere regions that performed high-level language computations during naturalistic sentence listening were also significantly activated during a non-language executive task. Although it has previously been shown that 'Multiple Demand'

regions become functionally engaged during more difficult than less difficult language tasks in healthy participants (Fedorenko et al., 2013), this study supports the view that domain-general executive regions can be functionally involved during language in healthy participants even during a naturalistic language task. Furthermore, rather than all domain-general regions being upregulated in PSA relative to controls (and this upregulation serving to compensate following damage to support language performance in PSA), this study demonstrates that there was in fact less language information processing in domain-general executive regions in PSA relative to controls, and the degree of functional disengagement of such domain-general executive regions correlates with the language impairment post-stroke via diaschisis.

These results change our interpretation of the role that domain-general executive regions perform during language in health and post-stroke aphasia. They suggest that domain-general executive regions might perform a core function during language in health that becomes necessary to maintain normal language performance in PSA. In keeping with this, a necessary role for domain-general executive regions during language in health was suggested by a previous lesion and functional imaging study which concluded that damage to the right inferior frontal sulcus results in comprehension impairment post-stroke because of this region's involvement in domain-general working memory (Gajardo-Vidal et al., 2018). The current study extends this by suggesting that the role of domain-general executive regions during language does not relate solely to 'executive' tasks and thus their involvement during language in PSA does not relate solely to increased requirements for cognitive control following damage to domain-specific language regions and increased task difficulty. Although 'Multiple Demand' regions are typically associated with executive control processes such as attention and strategy selection (Fedorenko et al., 2013), this study shows that domain-general executive regions process high-level language information during a naturalistic sentence comprehension task with minimal executive demands in health.

This poses the question as to what function such domain-general regions serve, and what is the nature of the computations they perform, during language. One proposal suggests that 'Multiple Demand' cortex contributes to goal-directed behaviour, in part, by holding a representation of the information that is currently being attended to which biases processing to attended or task-relevant information in other brain regions (Duncan, 2010). In this framework, a change to attended information should produce an update to the frontoparietal representation (Hon et al., 2006). Indeed, a previous study in healthy individuals instructed to watch a visual stream of words demonstrated that 'Multiple Demand' regions are activated when the attended sensory information (word) changed in the absence of any explicit task and thus no behaviour to perform (Hon et al.,

2006). This study goes further by demonstrating that changes to the frontoparietal representation of currently attended information can be produced not only by low-level changes in simple stimuli but also by high-level linguistic changes in the semantic predictability of the last word of a sentence. It is possible that the ability of 'Multiple-Demand' regions to process and/or detect high-level semantic changes to sentences is an important aspect of language processing that correlates with language performance post-stroke.

Chapter 5

The multidimensional nature of aphasia recovery post stroke

Manuscript accepted for publication at Brain

Abstract

Language is not a single function but instead results from interactions between neural representations and computations that can be damaged independently of each other. Although there is now clear evidence that the language profile in post-stroke aphasia reflects graded variations along multiple underlying dimensions ('components'), it is still entirely unknown if these distinct language components have different recovery trajectories and rely on the same, or different, neural regions during aphasia recovery. Accordingly, this study examined whether language components in the subacute stage: a) mirror those observed in the chronic stage; b) recover together in a homogenous manner; and c) have recovery trajectories that relate to changing activation in distinct or overlapping underlying brain regions. We analysed longitudinal data from 26 individuals with aphasia following left hemispheric infarct who underwent functional magnetic resonance imaging and behavioural testing at ~2 weeks and ~4 months post-stroke. The language profiles in early post-stroke aphasia reflected three orthogonal principal components consisting of fluency, semantic/executive function and phonology. These components did not recover in a singular, homogenous manner; rather, their longitudinal trajectories were uncorrelated, suggesting that aphasia recovery is heterogenous and multidimensional. Mean regional brain activation during overt speech production in unlesioned areas was compared with patient scores on the three principal components of language at both the early and late timepoints. In addition, the change in brain activation over time was compared with the change on each of the principal component scores, both before and after controlling for baseline scores. We found that different language components were associated with changing activation in multiple, non-overlapping bilateral brain regions during aphasia recovery. Specifically, fluency recovery was associated with increasing activation in bilateral middle frontal gyri and right temporooccipital middle temporal gyrus; semantic/executive recovery was associated with reducing activation in bilateral anterior temporal lobes; while phonology recovery was associated with reducing activation in bilateral precentral gyri, dorso-medial frontal poles, and the precuneus. Overlapping clusters in the ventromedial prefrontal cortex were positively associated with fluency recovery but negatively associated with semantic/executive and phonology recovery. This combination of detailed behavioural and fMRI data provides novel insights into the neural basis of aphasia recovery. Since different aspects of language seem to rely on different neural regions for recovery, treatment strategies that target the same neural region in all stroke survivors with aphasia might be entirely ineffective or even impair recovery, depending on the specific language profile of each individual patient.

Introduction

Aphasia affects at least one third of the more than 10 million new stroke cases globally each year (Benjamin et al., 2017; Engelter et al., 2006). Despite many stroke survivors exhibiting spontaneous partial language recovery (Hartman, 1981), difficulties persist into the chronic phase in at least 40% of initially aphasic patients (Wade et al., 1986). There is, therefore, an urgent need to understand the mechanisms involved in language recovery so that we can develop targeted treatments for this costly and debilitating condition (Ellis et al., 2012; Stefaniak et al., 2020; Tsouli et al., 2009).

Functional neuroimaging can provide empirical evidence regarding the patterns of neural activity that are associated with language performance post-stroke (Blank et al., 2003; Geranmayeh et al., 2016; Leff et al., 2002; Sharp et al., 2010; Stockert et al., 2020), which in turn might help to adjudicate between which underlying recovery mechanisms occur *in vivo* (Stefaniak et al., 2020). Language recovery occurs most rapidly during the first few months post-stroke (Pedersen et al., 1995), suggesting that imaging correlates of language performance in the subacute phase would provide the most information regarding recovery mechanisms. Despite the clear need for investigation of this early period post stroke, a recent formal meta-analysis of the fMRI aphasia literature found that studies were dominated by examination of very chronic cases (the median time post-onset across the literature was 38 months) and only a handful of studies provided longitudinal data albeit often in small samples (typical N<10) with minimal behavioural data (Stefaniak et al., 2021), crucial for understanding how behavioural variation relates to the fMRI results. The relative lack of studies, small sample sizes and minimal behavioural data are unsurprising given the considerable logistic challenges involved in recruiting, assessing and scanning subacute stroke patients. However, longitudinal studies are a powerful approach for exploring the neural bases of recovery (because the different starting points and inter-participant variations are controlled) and, particularly, for exploring whether language network changes observed in the chronic phase occur immediately or over time. Accordingly, this study utilised detailed behavioural data and task-based fMRI in a larger post-stroke aphasia (PSA) group at both early (2 weeks) and later (4 months) time points.

Amongst the handful of acute-to-chronic longitudinal PSA studies, several have associated better language recovery with increasing activation in regions of both hemispheres that are known to be part of the premorbid language network, including the left superior temporal gyrus (STG) (Cardebat et al., 2003; Heiss et al., 1999), left posterior temporal lobe (Stockert et al., 2020; van

Oers et al., 2018), left inferior frontal gyrus (IFG) (Stockert et al., 2020), right anterior temporal lobe (ATL) (Stockert et al., 2020), right STG (Cardebat et al., 2003; Nenert et al., 2018), right posterior temporal lobe (van Oers et al., 2018) and right IFG (Mattioli et al., 2014). Investigations have also associated better language recovery with increasing activation in domain-general executive regions (Geranmayeh et al., 2014), including the medial prefrontal cortex (Geranmayeh et al., 2017; Saur et al., 2006), left IFG pars opercularis and anterior insula (Stockert et al., 2020), left dorsolateral prefrontal cortex (Nenert et al., 2018), right anterior insula (Saur et al., 2006) and right dorsolateral prefrontal cortex (Nenert et al., 2018; Stockert et al., 2020). Regions associated with recovery in these studies were often activated in controls performing the same language task (Cardebat et al., 2003; Mattioli et al., 2014; Nenert et al., 2018; Saur et al., 2006; Stockert et al., 2020), consistent with the variable neuro-displacement principle, in which recovery follows from upregulation of the dynamic, spare capacity found in both undamaged regions of pre-morbid language networks and domain-general executive regions during subacute aphasia recovery (Stefaniak et al., 2020).

As noted above, a second major limitation of previous longitudinal studies of subacute PSA is the limited variety and depth of behavioural data. Typically, neuroimaging data are either correlated with a single behavioural score (Cardebat et al., 2003; Geranmayeh et al., 2017; Heiss et al., 1999; Radman et al., 2016; Saur et al., 2006; Stockert et al., 2020), or with a few language scores that cover a limited segment of the full language profile (Mattioli et al., 2014; Nenert et al., 2018; van Oers et al., 2018). However, language is not a single, homogenous cognitive function but is instead subserved by interactions between more general cognitive computations that can be damaged independently of each other (Mementi et al., 2011; Patterson & Lambon Ralph, 1999), thereby contributing to the multidimensional profile of chronic post-stroke aphasia (R. Alyahya et al., 2020; Butler et al., 2014; Halai et al., 2017; Mirman et al., 2015; Schumacher et al., 2019). In a cross-sectional study, Kummerer et al. used a data-driven decomposition technique, principal component analysis (PCA), to investigate the existence of distinct domains of language performance in subacute PSA and identified two components, representing ‘comprehension’ and ‘repetition’, that were associated with the degree of damage to the ventral and dorsal language pathways, respectively (Kummerer et al., 2013). Although there is now clear evidence that both subacute and chronic PSA reflects graded multidimensional variations, to date longitudinal studies have not investigated whether these distinct underlying components of language have different recovery trajectories, and if so, whether they are associated with changing activation in different or overlapping neural regions. Accordingly, this was a key target for the present study.

This study analysed data from one of the largest longitudinal functional neuroimaging studies in subacute post-stroke aphasia to date (Geranmayeh et al., 2017). The first question we addressed was whether the subacute language profile can be dissociated into orthogonal principal components that mirror those observed in the chronic stage, i.e. phonology (Butler et al., 2014), semantics (Butler et al., 2014), fluency (Halai et al., 2017) and executive function (Schumacher et al., 2019). The second question was whether such distinct underlying components recover in a homogenous manner, which we would expect if language recovery occurred uniformly along a single dimension, or whether the distinct behavioural components recover in a manner that is uncorrelated, which would suggest that aphasia recovery is multidimensional. The third question was whether different language components are associated with changing activation in distinct or overlapping neural regions during aphasia recovery. This has significant clinical implications; if different language components rely on different neural regions for recovery, then targeting the same neural region in all aphasic stroke survivors might be ineffective or even impair recovery, depending on the specific language profile of each individual patient.

In this study, we also examined a potentially crucial methodological issue for the first time. Perhaps unsurprisingly and unavoidably given the significant clinical challenges associated with the phase immediately after a stroke, the handful of studies that have examined acute PSA have recruited patients with mild-moderate aphasia. However, patients with mild aphasia in the subacute phase tend to recover well, resulting in their data at the chronic time point approaching ceiling and having reduced variance (Cardebat et al., 2003; Geranmayeh et al., 2017; Nenert et al., 2018; Stockert et al., 2020). This change in the shape of the data over time has important statistical consequences when examining the behavioural or imaging correlates of performance change. Recent work on the "proportional recovery rule" has demonstrated that, when chronic scores are less variable than at baseline, baseline scores inevitably become negatively correlated with subsequent performance change, even in the complete absence of any relationship between baseline and chronic scores (Hope et al., 2019). To our knowledge, no existing longitudinal imaging study has adequately accounted for baseline language performance during whole-brain image analysis. Consequently, regions identified during whole-brain image analysis in which activation change was associated with language recovery could be 'false positives', confounded by a true association between activation and baseline performance that is sufficient to account for the observed association with recovery. Additionally, there could be 'false negatives', in which a true association between activation and language recovery has been masked by varying baseline performance. This issue was examined formally in the current study.

Materials and methods

Participants

Twenty-seven individuals with a left hemisphere infarct who reported language difficulties at the onset of stroke, underwent a battery of language assessments and MRI scans at 2 weeks (Timepoint 1, T1) and 4 months (Timepoint 2, T2) post-stroke. Patients were premorbidly fluent in English and did not have a previous history of stroke resulting in aphasia or other neurological condition. 26 patients had a full battery of language test data available and were included in this analysis. Demographic and clinical variables for the PSA group are shown in Supplementary Table S5.1. 24 right-handed, fluent English speaking controls without history of neurological impairment were recruited for comparison, 22 of whom underwent two MRI scans over the same inter-scan interval as the patients. The 26 patients and 22 controls were age-matched (mean 59.0 [Standard Deviation 11.0] years in PSA vs. 57.6 [Standard Deviation 10.7] in controls; *t*-test, t_{46}=0.45, p=0.66). Informed consent was obtained from all participants according to the Declaration of Helsinki under approval by the National Research Ethics Service Committee. These data were used for previous publications which focused on neuroimaging correlates for a single language measure (Geranmayeh et al., 2017; Geranmayeh, Leech, et al., 2015; Geranmayeh et al., 2016; Geranmayeh et al., 2014). The current study provides entirely new results from extensive, novel analyses of the entire, combined behavioural and neuroimaging data spanning the two time points. In doing so, we were able to address the key study questions set out in the Introduction.

Neuropsychological tests

At both T1 and T2, patients completed subsections of the Comprehensive Aphasia Test (CAT) (Swinburn et al., 2005) yielding scores on 'Spoken Picture Description', 'Fluency', 'Spoken Comprehension', 'Written Comprehension', 'Repetition', 'Object Naming', 'Reading' and 'Cognitive'. They also underwent a brief version of the Raven's Progressive Matrix Test (Arthur & Day, 1994), and a quantitative analysis of connected speech production based on their narrative recollection of the Cinderella story yielding 'Information Carrying Words (ICWs) per second', 'Syllables per second', 'total ICWs', and a 'Narrative Aphasia Score' (NAS) (see Supplementary Methods) (Rochon et al., 2000; Saffran et al., 1989). In addition, several behavioural measures were obtained from their performance during the fMRI scan, including: Inverse Efficiency Score (IES) (mean reaction time in seconds/proportion correct) from the 'Decision' task (see below for

task details); 'Appropriate ICWs minus Inappropriate ICWs' (Swinburn et al., 2005) from the 'Speech' task (see below for task details); and 'Syllable Rate' from the 'Speech' task. This yielded 16 neuropsychological measures per patient at both T1 and T2. A subset of these neuropsychological measures were available in the controls at both time points. Neuropsychological test scores are shown in Supplementary Table S5.2 and S5.3.

Statistical analysis

We used varimax-rotated principal component analysis (PCA) to reduce the patients' scores across all 16 neuropsychological tests, to a smaller number of principal components (PCs), taken to be estimates of underlying cognitive dimensions. This data-driven decomposition technique has previously demonstrated the existence of orthogonal cognitive domains underlying observed variation in neuropsychological performance in chronic post-stroke aphasia (R. S. W. Alyahya et al., 2020; Butler et al., 2014; Halai et al., 2017; Mirman et al., 2015). Varimax-rotated PCA was performed on the 16 neuropsychological scores of patients, separately at T1 and T2. We used SPSS version 25 for the above statistical analyses. P-values are two-tailed and reported in their uncorrected form; alpha was set at 0.05, with statistical significance determined after applying a Bonferroni correction to the threshold (i.e. dividing 0.05 by the number of comparisons).

Lesion overlap map

The lesion overlap map of the PSA subgroup is shown in Fig. 5.1; it encompasses most of the left hemisphere, middle cerebral artery territory including subcortical white matter (Phan et al., 2005; Zhao et al., 2020).

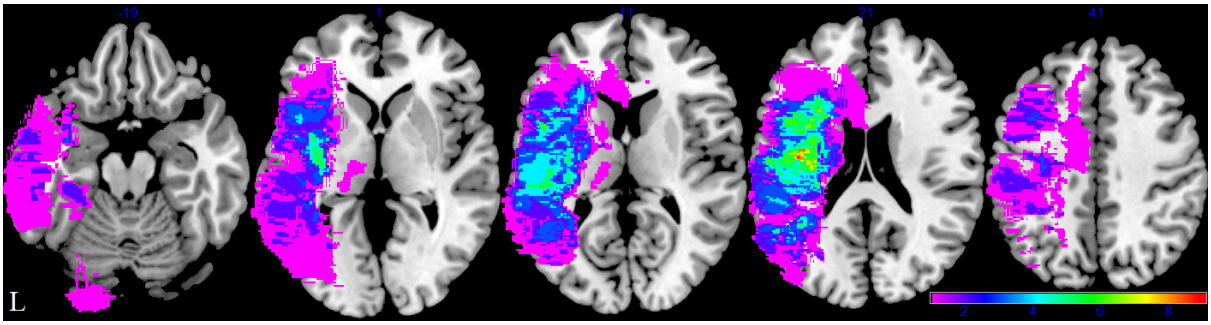


Figure 5.1: Lesion overlap map. Lesion overlap map for the 26 patient post-stroke aphasia subgroup in Montreal Neurological Institute space and in neurological convention using Timepoint 1 scans (2 weeks post-stroke). Colour bar represents overlap number between 1 and 9. The numbers refer to the Montreal Neurological Institute coordinate space (Montreal Neurological Institute) in the Z plane. L = Left.

Functional MRI design

For details of MRI data acquisition see Supplementary Methods. A sparse fMRI design was used to minimise motion artefact by acquiring EPI volumes in between periods of speech production (Gracco et al., 2005; Hall et al., 1999). Each scanning session contained three fMRI runs. Each run contained 20 'Speech' trials, 16 'Count' trials, 16 'Decision' trials and 15 'Rest' trials. Each 'Speech' trial required participants to overtly define an object shown as a coloured picture; the speech produced was recorded with an Optacoustics FOMRI-III microphone, transcribed and analysed to yield the behavioural measures described previously. 'Count' trials required participants to count from one to seven at a rate of 1/s; 'Decision' trials required participants to press a button when presented with a blue square but inhibit their response when presented with an orange circle; and 'Rest' trials required participants to view a fixation cross. Trials were presented in blocks of two or four of the same type, with each of the four 'Decision' blocks per run being preceded by a 10s instruction page. Each trial lasted 10s; the task was performed during the first 7s before being terminated by a fixation cross, with the EPI fMRI volume being acquired 1s later over 2s. Thus, one EPI volume was acquired per trial, and 71 volumes were acquired over each run.

Functional MRI first-level analysis

Each subject's EPI volumes were realigned, brain-extracted, smoothed, coregistered and normalised using FMRIB Software Library (www.fmrib.ox.ac.uk/fsl) (Jenkinson et al., 2002; Jenkinson et al., 2012; Smith, 2002) and Statistical Parametric Mapping (SPM) 12 (Wellcome Centre for Human Neuroimaging, London UK; www.fil.ion.ucl.ac.uk/spm/) (Seghier et al., 2008). See Supplementary Methods for more detail. The preprocessed data was then entered into a fixed-effect first-level analysis using the general linear model in which each 10s trial was modelled as an epoch convolved with the canonical haemodynamic response function (HRF) (Friston et al., 1995). Regressors were resampled 9s into each 10 second trial, which was the half-way point of each EPI volume's acquisition in this sparse scanning design. Six rigid-body motion parameters were included as regressors of no-interest. Each subject's three runs were modelled as separate sessions in the same first-level design matrix.

The contrast of interest was the summation of 'Speech and Count' over 'Rest', which represents the combined activation during overt speech production over the course of the scan. This contrast theoretically utilises all aspects of language processing, including connected speech production (fluency) (R. Alyahya et al., 2020), speech sound processing (phonology) (Schwartz et al., 2012),

and the controlled retrieval of information related to an object (semantics) (Lambon Ralph et al., 2017), the latter of which should rely on regions that are commonly used during semantic processing in both production and comprehension tasks (Hodges et al., 1992; Pobric et al., 2007). This single contrast was therefore likely to identify neural correlates of multiple language dimensions, while maximising the amount of imaging data included and minimising the multiple comparisons that would have occurred if multiple separate contrasts were used.

ImCalc was used to create a 'mean activation image' from the mean of each subject's first and second timepoint 'Speech+Count>Rest' contrast images. These 'mean activation' images were entered into a second-level analysis assessing the main effect of group (patients vs. controls) on regional activation during speech production (outlined below). ImCalc was also used to subtract each subject's first timepoint 'Speech+Count>Rest' contrast image from the corresponding second timepoint contrast image to obtain an 'activation change' image representing the change in parameter estimate over time. These 'activation change' images were entered into second-level analyses with language change, i.e. the change over time of each of the three PC scores, as covariates (outlined below).

Functional MRI second-level analysis

First-level 'Speech+Count>Rest' contrast estimates from each subject at each voxel were entered into group-level random-effects analyses in SPM12. Statistical thresholding used a voxel-wise cluster forming threshold of $p<0.005$ (uncorrected) and a cluster-level threshold of $p<0.05$ after familywise error correction. In order to minimise the confounding effect of variable lesion morphology, analyses were restricted to grey matter voxels in which no patient had a lesion. This resulted in a large proportion of the left hemisphere being excluded.

To identify whether regional activation during speech production varied between participant groups (patients vs. controls) and timepoints (T1 vs. T2), we performed a second-level mixed-design ANOVA using SPM's 'flexible factorial' option with first-level 'Speech+Count>Rest' contrast estimate as the dependent variable, 'timepoint' as the within-subjects factor and 'participant group' as the between-subjects factor. This assessed for a 'participant group*timepoint' interaction and a main effect of timepoint (T1 vs. T2); since the error term for this model was within-subjects, a main effect of group was assessed using an independent sample *t*-test comparing the 'mean activation images' of patients vs. controls.

To identify regions in which activation associated with language performance, patient group 'Speech+Count>Rest' contrast estimates were entered into second-level analyses with the three

neuropsychological principal components (PCs) included as regressors of interest. T1 brain activation was related to T1 PC1, T1 PC2 and T1 PC3 scores in three separate analyses; 'activation change' (obtained by subtracting each participant's contrast image at T1 from T2) was associated with change over time in scores of each of the three PCs (PC1 change, PC2 change and PC3 change, obtained by subtracting each participant's PC score at T1 from T2) in three separate analyses; and T2 activation was associated with T2 PC1, T2 PC2 and T2 PC3 scores in three separate analyses. In addition, we wanted to identify regions in which activation change was associated with language performance change after controlling for baseline severity of the language deficit. We therefore repeated second-level analyses in which activation change was associated with PC1 change, PC2 change and PC3 change, including T1 PC1, T1 PC2 and T1 PC3 scores as regressors of no interest in their respective separate design matrices.

We extracted the mean parameter estimate from each cluster identified as being associated with language performance in the above analyses using MarsBaR 0.44 (Brett et al., 2002). Mean parameter estimate here represents either mean activation during 'Speech+Count>Rest' at T1 or T2 or the change in mean activation during 'Speech+Count>Rest' between T1 and T2. To account for the relationship between a given cluster-derived parameter estimate and language scores above and beyond that explained by cofounding factors, we used robust regression implemented in Matlab2018a to construct several analyses per cluster parameter estimate with language PC score as the dependent variable. The first model included mean parameter estimate alone. For clusters assessing activation change between T1 and T2, three additional models were constructed. The first additional model included mean activation change as well as T1 PC score. If mean activation change was significant in the first additional model, a second additional model added T1 PC*mean parameter estimate to check for a significant interaction, and a third additional model included the first or second additional model (depending on whether the interaction was significant) as well as lesion volume, years of education and age. Regression coefficients are reported as unstandardised.

Anatomical labels were defined using the Harvard-Oxford atlas for cortical regions (Desikan et al., 2006) and the Automated Anatomical Labelling atlas for subcortical regions (Tzourio-Mazoyer et al., 2002).

See the online Supplementary Methods for additional details regarding participants, neuropsychological testing, statistical analyses, fMRI design and processing.

Data availability

The data of this study are available upon reasonable request.

Results

Neuropsychological tests

At the group level, patients with aphasia were significantly impaired relative to controls on virtually all tests at T1 (Supplementary Table S5.4). Paired statistical tests confirmed that patients performed significantly better at 4 months (T2) than 2 weeks (T1) post-stroke on all 16 measures, except for 'Decision Task IES' and 'Speech Task Syllable Rate' (Supplementary Table S5.5) where their performance remained the same. Consequently, patient performance at T2 started to approach control levels, with 'CAT Cognitive' and 'Decision Task IES' no longer being significantly impaired compared to the controls (Supplementary Table S5.6). Each of the 16 neuropsychological tests had a smaller interquartile range (IQR) at T2 than T1, suggesting less within group variance (Supplementary Table S5.5).

Principal Component Analysis of the neuropsychological scores

We performed varimax-rotated PCA on the correlation matrix of neuropsychological test scores of patients at T1. The Kaiser-Meyer-Olkin value was 0.79, indicating adequate sampling (Kaiser, 1974). Bartlett's test of sphericity was significant (approximate $\chi^2_{120}=465$, $p=4x10^{-42}$), suggesting these data could be factorised. Three rotated principal components (PCs) with eigenvalues greater than 1 explaining 30.45% (PC1), 25.37% (PC2) and 22.84% (PC3) of the variance were obtained (Table 5.1). The scree plot had an inflection point after three PCs (Supplementary Figure 5.1), confirming that 3 components should be retained (Cattell, 1966).

PC1 was primarily loaded onto by measures of connected speech production obtained from analysis of the Cinderella story ('Cinderella ICWs Per Second', 'Cinderella Syllables Per Second', 'Cinderella Total ICWs' and 'Cinderella NAS') as well as 'CAT Spoken Picture Description' and 'CAT Fluency'. Thus, PC1 was interpreted as representing fluency of connected speech (Table 5.1). PC2 was loaded onto primarily by 'CAT Spoken Comprehension', 'CAT Written Comprehension', 'CAT Cognitive', 'Ravens' and in-scanner 'Decision Task IES' (lower scores representing better performance thus 'Decision Task IES' loaded negatively onto PC2). Thus, PC2 was interpreted as representing semantic/executive performance (Table 5.1). PC3 was loaded onto primarily by performance on 'CAT Repetition', 'CAT Object Naming', 'CAT Reading' and in-scanner 'Speech Task Syllable Rate'. Thus, PC3 was interpreted as representing

phonological ability (Table 5.1). These components representing fluency (PC1), semantic/executive function (PC2) and phonology (PC3) strongly resemble the principal components that were previously obtained using independent data from chronic post-stroke aphasia (Butler et al., 2014; Halai et al., 2017; Mirman et al., 2015), as well as in one previous analysis of acute PSA (Kummerer et al., 2013). This PCA structure has been formally shown to be highly stable in chronic aphasia (Halai, Woollams, et al., 2020) and has been replicated across different research groups.

Neuropsychological test	Component loadings		
	PC 1	**PC 2**	**PC3**
Cinderella ICWs Per Second	**0.93**	0.08	0.24
Cinderella Syllables Per Second	**0.89**	0.12	0.23
Cinderella Total ICWs	**0.86**	0.18	0.13
Cinderella NAS	**0.80**	0.27	0.44
CAT Spoken Picture Description	**0.73**	0.21	0.46
CAT Fluency	**0.64**	0.43	0.48
CAT Spoken Comprehension	0.17	**0.91**	0.25
CAT Written Comprehension	0.24	**0.83**	0.33
Decision Task IES	0.10	**-0.76**	-0.10
Ravens	0.42	**0.69**	0.14
CAT Cognitive	0.36	**0.61**	0.29
Speech Task Appropriate Minus Inappropriate ICWs	0.42	0.53	0.51
CAT Repetition	0.22	0.16	**0.89**
CAT Object Naming	0.29	0.44	**0.78**
CAT Reading	0.27	0.51	**0.72**
Speech Task Syllable Rate	0.35	0.15	**0.67**

Table 5.1: Component matrix of neuropsychological scores from patients with post-stroke aphasia at Timepoint 1. Varimax rotated principal component analysis was performed on the neuropsychological scores of patients with post-stroke aphasia at Timepoint 1 (2 weeks post-stroke). The loading of each score onto each rotated principal component is shown. Variables with major loadings (defined as >0.60) are in bold. Abbreviations: CAT = Comprehensive Aphasia Test; ICW = Information Carrying Word; IES = Inverse Efficiency Score; NAS = Narrative Aphasia Score.

We also performed varimax-rotated PCA on the correlation matrix of neuropsychological test scores of patients at T2. The Kaiser-Meyer-Olkin value was 0.46, indicating inadequate sampling (Kaiser, 1974). Since each of the 16 neuropsychological tests had a smaller interquartile range

(IQR) at T2 than T1 (Supplementary Table S5.5), it seems likely that reduced test score variation at T2 caused the T2 data to be inadequate for PCA (which requires sufficient variation in the measures to generate any form of structure).

To enable direct comparisons to be made between T1 and T2, the patients' scores need to be in the same PCA 'space'. As the T2 data were not appropriate for PCA reduction (see above), the neuropsychological scores at T2 were back-projected into the 'T1 PCA space' (see Supplementary Methods). Individual PC scores are shown in Supplementary Table S5.7. PC1 scores were significantly better at T2 than T1 (median -0.02 [IQR 1.20] at T1 vs 0.62 [0.96] at T2; Wilcoxon signed-rank test, $z=3.67$, $p<0.0005$). Similarly, PC2 scores were significantly better at T2 than T1 (median 0.43 [IQR 1.06] at T1 vs 0.51 [0.64] at T2; Wilcoxon signed-rank test, $z=3.01$, $p=0.003$). However, PC3 scores were not significantly better at T2 than T1 (median 0.21 [IQR 1.07] at T1 vs 0.26 [0.59] at T2; Wilcoxon signed-rank test, $z=0.88$, $p=0.38$). Like the raw neuropsychological scores, the three PCs had a smaller IQR at T2 compared to T1. Following the formal demonstrations by Hope *et al.*, we therefore expected T1 PC scores to be negatively correlated with the change in performance over time, even if there were no underlying correlation between scores at T1 and T2 (Hope et al., 2019). Indeed, T1 PC1 was significantly negatively correlated with PC1 change (Spearman's rho=-0.55, $p=0.003$); T1 PC2 was significantly negatively correlated with PC2 change (Spearman's rho=-0.57, $p=0.002$); and T1 PC3 was significantly negatively correlated with PC3 change (Spearman's rho=-0.81, $p=5x10^{-7}$). It was therefore important to control for baseline PC score when considering the neural correlates of performance change. If one looked at the change measures alone, the interpretation of them would be ambiguous as the regions might reflect change or simply the baseline performance itself (see below).

If aphasia recovery occurred in a unidimensional manner, one would expect the different 'PC change' scores to be strongly, positively correlated. However, there was no significant correlation between PC1 change and PC2 change (Spearman's rho=-0.18, $p=0.39$) nor between PC2 change and PC3 change (Spearman's rho=0.20, $p=0.32$), while PC1 change was significantly negatively correlated with PC3 change (Spearman's rho=-0.62, $p=0.001$). These results were unchanged after partialling out T1PC1, T1PC2 and T1PC3 scores from each of the previous correlations (PC1 change vs PC2 change, Spearman's rho=-0.47, $p=0.03$; PC2 change vs PC3 change, Spearman's rho=0.19, $p=0.39$; PC1 change vs PC3 change, Spearman's rho=-0.53, $p=0.009$). These results suggest that aphasia recovery is multidimensional and raises the intriguing

possibility that different components of language might recover in an anticorrelated manner, as better fluency recovery was associated with poorer phonological recovery in this sample.

Figure 5.2 contains three scatterplots depicting the 26 patients moving through PC1-PC2 space (Fig. 5.2A), PC1-PC3 space (Fig. 5.2B) and PC2-PC3 space (Fig. 5.2C) between 2 weeks and 4 months post-stroke. If there was a singular recovery process then the patients would move uniformly towards the upper, right-hand corner (towards control level performance). Instead, it is clear that the patients did not 'move' together through each PCA space but rather had different recovery trajectories (Fig. 5.2). For instance, several patients improve predominantly along one PC dimension while remaining static on the other PC dimension (shown in the scatterplots as moving vertically or horizontally), while some patients are already within the limits of 'normal' at T1 and remain in a similar position of PCA space at T2 (Fig. 5.2).

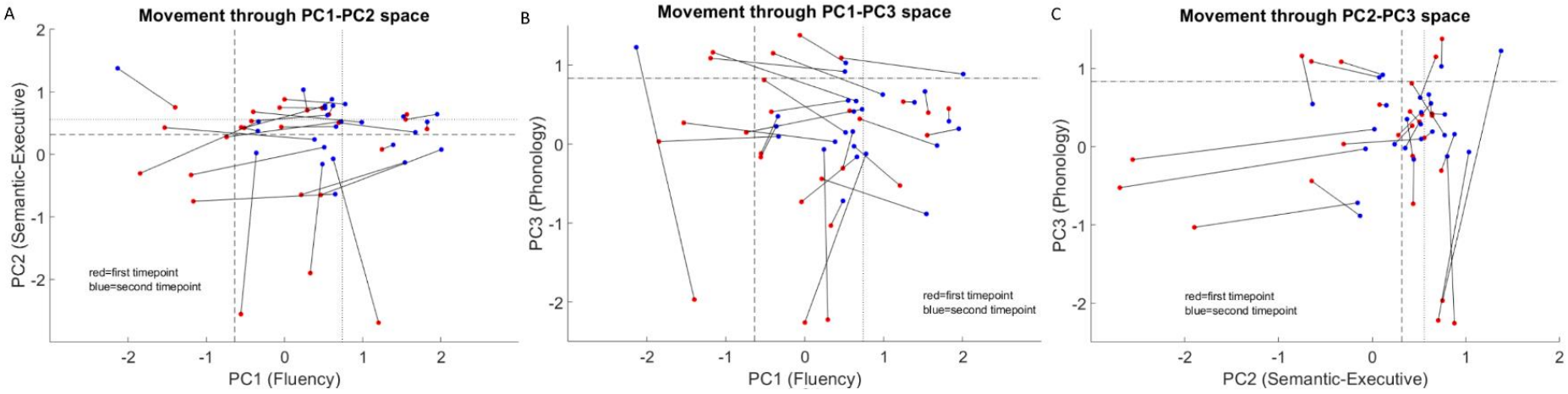


Figure 5.2: Movement through PCA space during recovery. Scatter plots depicting the 26 patients with post-stroke aphasia moving through Principal Component 1-Principal Component 2 space **(A)**, Principal Component 1-Principal Component 3 space **(B)** and Principal Component 2-Principal Component 3 space **(C)** between 2 weeks and 4 months post-stroke. Each line represents an individual patient with the red circle their Timepoint 1 performance (2 weeks post-stroke) and the blue circle their Timepoint 2 performance (4 months post-stroke). Dashed lines represent the 'lower bound of normal control performance'; dotted lines represent the 'mean Principal Component score'; X and Y axis show the component scores. Abbreviations: PC = Principal Component; PC1 = 'fluency' Principal Component; PC2 = 'semantic/executive' principal component; PC3 = 'phonology' principal component.

Regional activation during speech production

A one-sample *t*-test on the 'mean activation image', averaged across T1 and T2, of 'Speech+Count>Rest' in healthy controls identified significant bilateral activation throughout frontal cortex (precentral gyrus, supplementary motor cortex, superior frontal gyrus, middle frontal gyrus [MFG], IFG pars opercularis and left IFG pars triangularis), temporal cortex (Heschl's gyrus, planum temporale, planum polare, STG, middle temporal gyrus [MTG], temporal pole), parietal cortex (postcentral gyrus, superior parietal lobule, supramarginal gyrus), insular cortex, anterior cingulate cortex, putamen and thalamus (Fig. 5.3A, Supplementary Table S5.8). This demonstrates that multiple regions throughout both hemispheres are involved in language in health. The opposite contrast, 'Rest>Speech+Count', identified significant bilateral deactivation during speech in the frontal medial cortex, subcallosal cortex, frontal pole, angular gyrus, posterior cingulate gyrus and precuneus (Fig. 5.3B, Supplementary Table S5.8), i.e., core regions of the Default Mode Network (Margulies et al., 2016).

To identify whether regional activation during speech production varied between participant groups (patients vs. controls) and timepoints (T1 vs. T2), we performed a mixed-design ANOVA with first-level 'Speech+Count>Rest' contrast estimate as the dependent variable, 'timepoint' as the within-subjects factor and 'participant group' as the between-subjects factor. The main effect of timepoint (T1 vs. T2) was not significant, meaning that regional activation during speech production, collapsed across patients and controls, did not change significantly between 2 weeks and 4 months post-stroke. The 'participant group*timepoint' interaction was not significant, meaning that activation change between timepoints was not significantly different in patients compared to controls. The main effect of group was assessed using an independent samples *t*-test comparing the 'mean activation images', averaged across timepoints, of patients vs. controls. We did not identify any regions of significantly greater activation in patients than controls. We found significantly less activation in patients than controls in three clusters. The first cluster encompassed right posterior cingulate and temporo-parietal cortex including the precuneus, superior parietal lobule, angular gyrus, posterior supramarginal gyrus, STG, posterior MTG, posterior inferior temporal gyrus (ITG) and temporal pole (Fig. 5.3C, Supplementary Table S5.9); this cluster overlapped with regions activated in controls in the right planum temporale, anterior STG and temporal pole, and with regions deactivated in controls in the right precuneus and angular gyrus. The second cluster was in the left precuneus, posterior cingulate gyrus, superior parietal lobule, thalamus and caudate; while the third encompassed the right thalamus and pallidum (Fig. 5.3C, Supplementary Table S5.9).

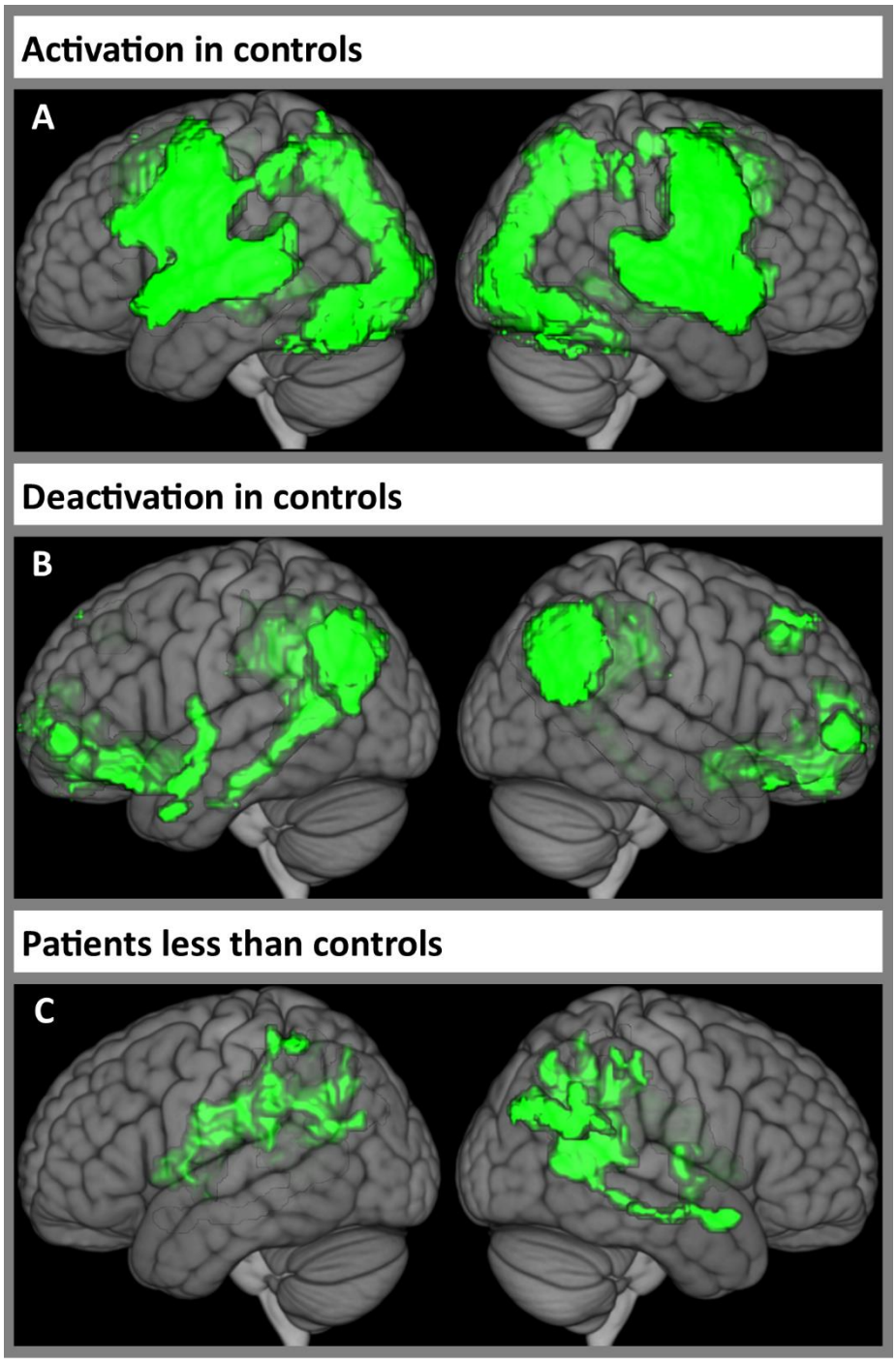


Figure 5.3: Regional activation during overt speech production. (A) Regions of significant activation during 'Speech+Count>Rest', averaged across 2 weeks and 4 months post-stroke, in healthy controls (Supplementary Table S5.8). **(B)** Regions of significant deactivation during 'Speech+Count>Rest', averaged across 2 weeks and 4 months post-stroke, in healthy controls (Supplementary Table S5.8). **(C)** Regions of significantly less activation in patients than controls, averaged across 2 weeks and 4 months post-stroke (Supplementary Table S5.9). The left and right columns show the left and right hemispheres, respectively. Statistical thresholding used a voxel-wise cluster forming threshold of $p<0.005$ (uncorrected) and a cluster-level threshold of $p<0.05$ after familywise error correction.

Activation positively associated with fluency at 2 weeks post-stroke

We did not identify any clusters in which T1 activation was associated with T1 'semantic/executive' (PC2) or 'phonology' (PC3) score. However, T1 activation in a cluster in the right posterior supramarginal gyrus, insular cortex and temporooccipital MTG was positively associated with T1 'fluency' (PC1) score (Fig. 5.4, Supplementary Table S5.10). Mean activation extracted from this cluster was significantly positively associated with T1 PC1 score (beta=0.54, $p=7.1x10^{-5}$), even after including lesion volume, years of education and age (beta=0.51, p=0.0004) (Fig. 5.4, Supplementary Table S5.11).

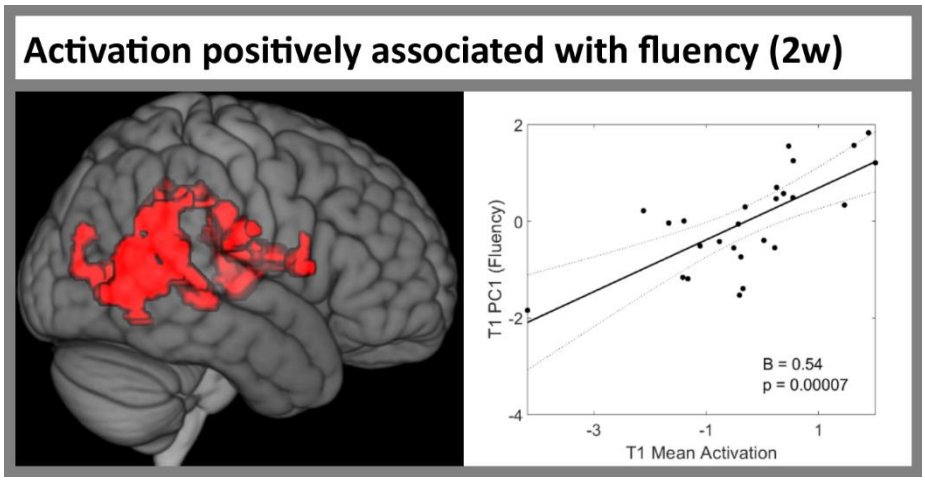


Figure 5.4: Region in which activation was positively associated with fluency at 2 weeks post-stroke. The left column shows the cluster from Supplementary Table S5.10 in which activation during 'Speech+Count>Rest' was positively associated with Principal Component 1 'fluency' score in patients with post-stroke aphasia at 2 weeks post-stroke. The right column scatter plot shows the significant, positive association between mean activation in this cluster (x-axis) and Principal Component 1 'fluency' score at 2 weeks post-stroke (y-axis) from the robust regression model in Supplementary Table S5.11. Statistical thresholding used a voxel-wise cluster forming threshold of $p<0.005$ (uncorrected) and a cluster-level threshold of $p<0.05$ after familywise error correction. Abbreviations: B = unstandardised regression coefficient; PC = Principal Component; PC1 = 'fluency' Principal Component; T1 = Timepoint 1 (2 weeks post-stroke).

Activation change positively associated with fluency improvement between 2 weeks-4 months post-stroke

Before including baseline 'fluency' score (PC1) in the mass univariate analysis, activation change in three clusters was significantly positively associated with PC1 improvement (Supplementary Table S5.12). The first cluster was centred in the left MFG (Fig. 5.5A); the second cluster was centred in the right temporooccipital MTG (Fig. 5.5B), in a location similar to the cluster associated with PC1 score at T1 (Fig. 5.4); and the third cluster was in the right MFG (Fig. 5.5C). Mean activation change extracted from each of these clusters was significantly positively associated with PC1 improvement both before and after including T1 PC1 score, and even after

including lesion volume, years of education and age (cluster 1 beta=0.29, p=0.0003; cluster 2 beta=0.16, p=0.04; cluster 3 beta=0.30, p=0.0001) (Fig. 5.5, Supplementary Table S5.13).

After including baseline 'fluency' (PC1) score in the mass univariate analysis, activation change in two clusters was significantly positively associated with better PC1 improvement (Supplementary Table S5.14). The first cluster was centred in the bilateral ventromedial prefrontal cortex (Fig. 5.5D). The second cluster was centred in the right temporooccipital MTG (Fig. 5.5E), in a location similar to the clusters associated with PC1 at T1 (Fig. 5.4) and with PC1 change before including T1 PC1 in the mass univariate analysis (Fig. 5.5B). Mean activation change extracted from both clusters was significantly positively associated with PC1 improvement both before and after including T1 PC1 score, and even after including lesion volume, years of education and age (cluster 1 beta=0.38, p=0.0001; cluster 2 beta=0.23, p=0.003) (Fig. 5.5, Supplementary Table S5.15).

We did not identify any clusters in which activation change was significantly negatively associated with 'fluency' PC1 change.

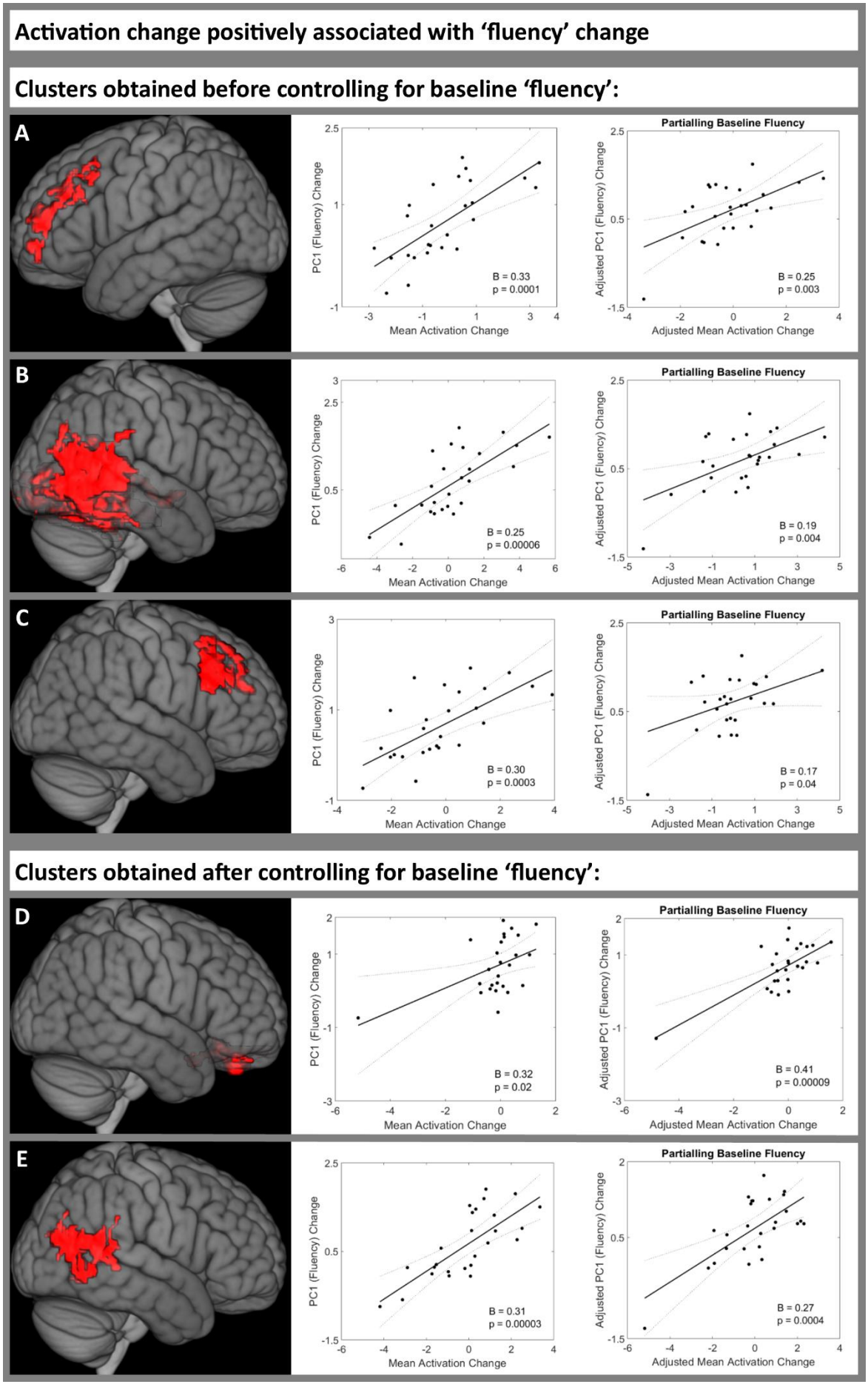


Figure 5.5: Regions in which increased activation was positively associated with fluency improvement between 2 weeks and 4 months post-stroke. Activation change was positively associated with Principal Component 1 'fluency' score change between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-

stroke) in patients with post-stroke aphasia. Three clusters were significant before controlling for baseline 'fluency' score **(A-C)** (Supplementary Table S5.12, S5.13), two clusters were significant after controlling for baseline 'fluency' score **(D-E)** (Supplementary Table S5.14, S5.15). The left column shows each cluster rendered on the Montreal Neurological Institute standard brain template. The middle column shows the association between the mean activation change of each cluster and Principal Component 1 'fluency' score change, using robust regression. The right column shows the association between the mean activation change of each cluster and Principal Component 1 'fluency' score change, controlling for baseline 'fluency' score, using robust regression. Statistical thresholding used a voxel-wise cluster forming threshold of $p<0.005$ (uncorrected) and a cluster-level threshold of $p<0.05$ after familywise error correction. Abbreviations: B = unstandardised regression coefficient; PC = Principal Component; PC1 = 'fluency' Principal Component; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post-stroke).

Activation change negatively associated with semantic/executive change between 2 weeks-4 months post-stroke

We did not identify any clusters in which activation change was significantly positively associated with PC2 change.

Before including baseline 'semantic/executive' (PC2) score in the mass univariate analysis, no clusters were significantly negatively associated with PC2 change. However, after including baseline PC2 score in the mass univariate analysis, activation change in two clusters was significantly negatively associated with PC2 change (Supplementary Table S5.16). The first cluster was in the left temporal pole, frontal medial cortex and frontal pole (Fig. 5.6A). The second cluster encompassed the right temporal pole, anterior MTG and posterior ITG (Fig. 5.6B). These two clusters therefore encompass core regions of the bilateral semantic network (Lambon Ralph et al., 2017). For both clusters, extracted mean activation change was not associated with PC2 change before including T1 PC2 score (cluster 1 beta=-0.01, p=0.84; cluster 2 beta=-0.03, p=0.79) (Fig. 5.6, Supplementary Table S5.17). However, after including T1 PC2 score, 'mean activation change' became significantly negatively associated with PC2 change in both clusters (cluster 1 beta=-0.22, $p=2.5\times10^{-7}$; cluster 2 beta=-0.28, $p=8.0\times10^{-5}$) (Fig. 5.6, Supplementary Table S5.17). This remained true after adding lesion volume, years of education and age (cluster 1's beta=-0.22, $p=3.0\times10^{-6}$; cluster 2's beta=-0.29, $p=5.2\times10^{-5}$) (Supplementary Table S5.17). These results demonstrate that including baseline language performance in the mass univariate analysis can identify novel areas that are associated with language change but would otherwise remain obscured.

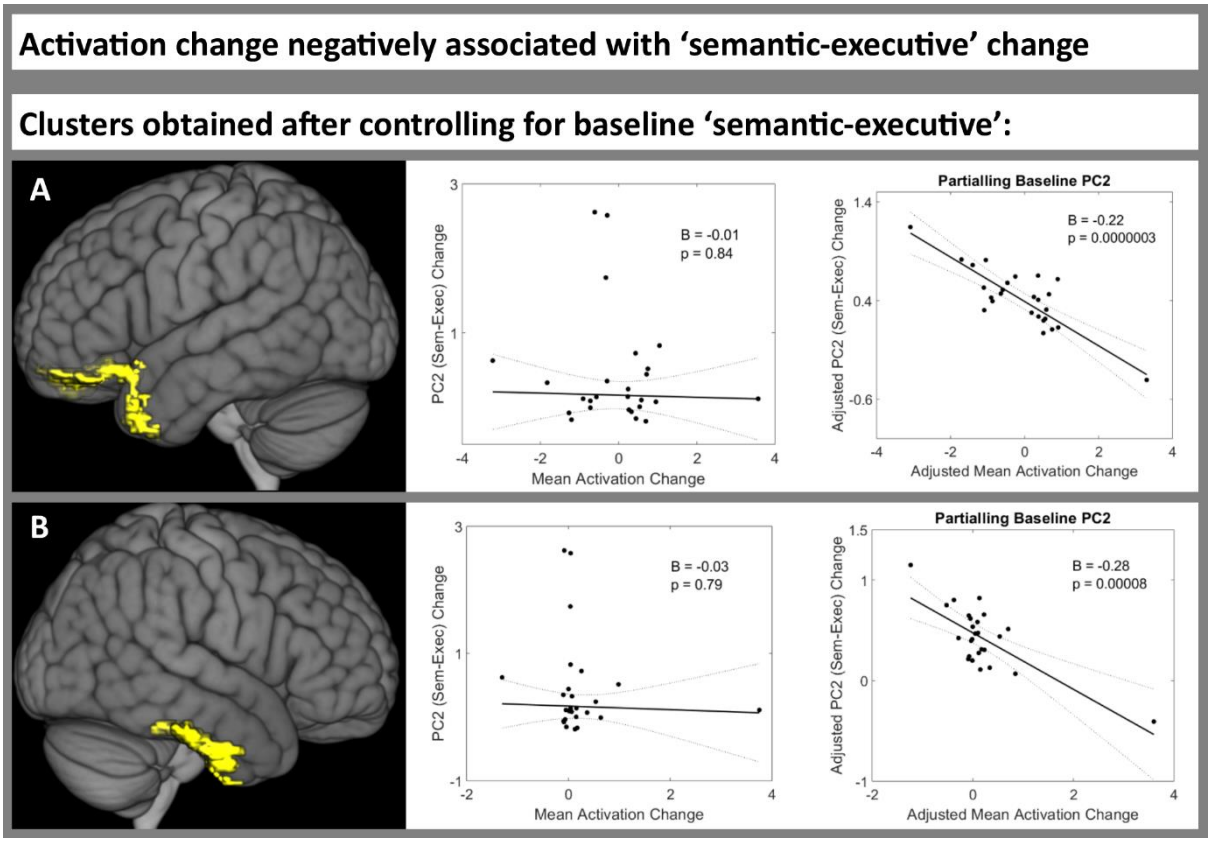


Figure 5.6: Regions in which increased activation was negatively associated with semantic/executive improvement between 2 weeks and 4 months post-stroke. Activation change was negatively associated with Principal Component 2 'semantic/executive' score change between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) in patients with post-stroke aphasia. No clusters were significant before controlling for baseline 'semantic/executive' score, but two clusters were significant after controlling for baseline 'semantic/executive' score **(A-B)** (Supplementary Table S5.16, S5.17). The left column shows each cluster rendered on the Montreal Neurological Institute standard brain template. The middle column shows the association between the mean activation change of each cluster and Principal Component 2 'semantic/executive' score change, using robust regression. The right column shows the association between the mean activation change of each cluster and Principal Component 2 'semantic/executive' score change, controlling for baseline 'semantic/executive' score, using robust regression. Statistical thresholding used a voxel-wise cluster forming threshold of $p<0.005$ (uncorrected) and a cluster-level threshold of $p<0.05$ after familywise error correction. Abbreviations: B = unstandardised regression coefficient; PC = Principal Component; PC2 = 'semantic/executive' principal component; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post stroke).

Activation change negatively associated with phonology change between 2 weeks-4 months post-stroke

We did not identify any clusters in which activation change was significantly positively associated with 'phonology' (PC3) change.

Before including baseline PC3 in the mass univariate analysis, activation change in three clusters was significantly negatively associated with PC3 change (Supplementary Table S5.18). The first

cluster affected the bilateral precentral gyrus, MFG and superior frontal gyrus (Fig. 5.7A). Mean activation change extracted from the first cluster was significantly negatively associated with PC3 change (beta=-0.27, $p=8.3x10^{-5}$); however, when including T1 PC3 score and 'T1 PC3*mean activation change' in the model, it transpired that there was a significant main effect of T1 PC3 (beta=-0.49, $p=6.9x10^{-5}$) and a significant interaction ('T1 PC3*mean activation change' beta=0.14, p=0.0009) such that 'mean activation change' was only negatively associated with PC3 change in patients with low PC3 scores at T1 (Fig. 5.7A, Supplementary Table S5.19). The second cluster was in bilateral ventromedial prefrontal cortex (Fig. 5.7B), similar to the clusters that were positively associated with PC1 change (Fig. 5.5D) and negatively associated with PC2 change (Fig. 5.6A). However, mean activation change extracted from cluster 2 was not significantly associated with PC3 change after including T1 PC3 (beta=-0.08, p=0.09) (Fig. 5.7B, Supplementary Table S5.19). The third cluster was in the left frontal pole (Fig. 5.7C), partially overlapping with the cluster positively associated with PC1 change (Fig. 5.5A). However, mean activation change extracted from cluster 3 was not significantly negatively associated with PC3 change after including T1PC3, lesion volume, years of education and age (beta=-0.11, p=0.08) (Fig. 5.7C, Supplementary Table S5.19).

After including baseline 'phonology' (PC3) score in the mass univariate analysis, activation change in three clusters was significantly negatively associated with PC3 change (Supplementary Table S5.20). The first cluster affected the bilateral frontal poles and midline superior frontal gyrus (Fig. 5.7D), more medially than the bilateral MFG clusters positively associated with PC1 change (Fig. 5.5A, 5.5C). Mean activation change extracted from cluster 1 was not associated with PC3 change before including T1 PC3 (beta=-0.05, p=0.50), but was negatively associated with PC3 change after including T1 PC3 (beta=-0.13, p=0.0006) and had a significant 'T1 PC3*mean activation change' interaction ('T1 PC3*mean activation change' beta=0.08, p=0.05) such that 'mean activation change' was only negatively associated with PC3 change in patients with low PC3 scores at T1 (Fig. 5.7D, Supplementary Table S5.21). The second cluster affected the ventromedial prefrontal cortex (Fig. 5.7E), similar to the clusters that were positively associated with PC1 change (Fig. 5.5D) and negatively associated with PC2 change (Fig. 5.6A). Mean activation change extracted from cluster 2 was negatively associated with PC3 change both before and after including T1 PC3 score, and even after including lesion volume, years of education and age (beta=-0.28, p=0.0001) (Fig. 5.7E, Supplementary Table S5.21). The third cluster was in the precuneus (Fig. 5.7F). Mean activation change extracted from cluster 3 was negatively associated with PC3 change before and after including T1 PC3 (model including T1

PC3 and interaction term, beta=-0.15, p=0.0002) and had a significant ‘T1 PC3*mean activation change’ interaction (‘T1 PC3*mean activation change’ beta=0.09, p=0.03) such that ‘mean activation change’ was only negatively associated with PC3 change in patients with low PC3 scores at T1 (Fig. 5.7F, Supplementary Table S5.21).

We did not identify any clusters in which T2 activation was associated with T2 PC1, PC2 or PC3 score.

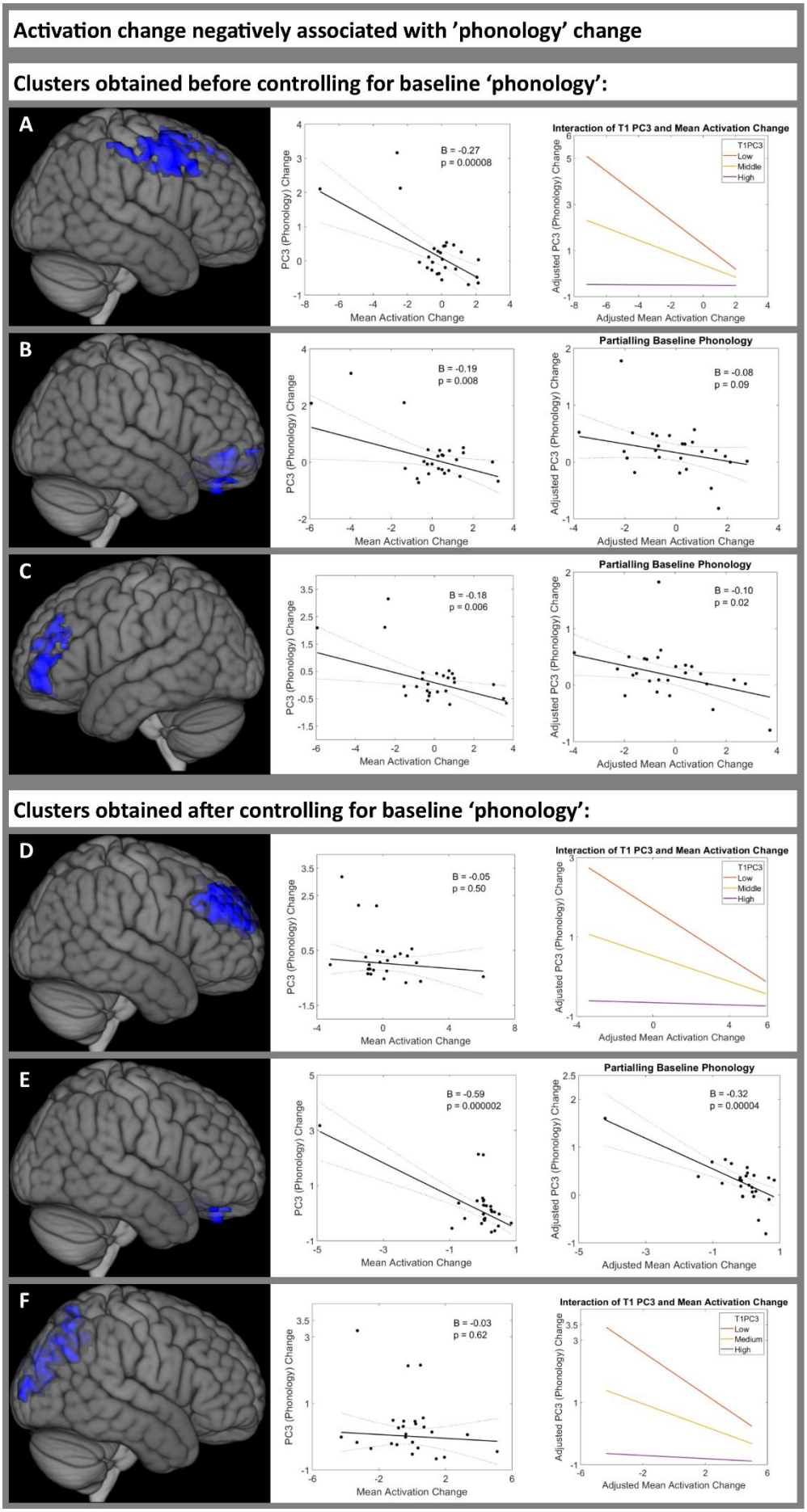


Figure 5.7: Regions in which increased activation was negatively associated with phonology improvement between 2 weeks and 4 months post-stroke. Activation change was negatively associated with Principal Component 3 'phonology' score change between Timepoint 1 (2 weeks) and Timepoint 2 (4

months post-stroke) in patients with post-stroke aphasia. Three clusters were significant before controlling for baseline 'phonology' score **(A-C)** (Supplementary Table S5.18, S5.19), three clusters were significant after controlling for baseline 'phonology' score **(D-F)** (Supplementary Table S5.20, S5.21). The left column shows each cluster rendered on the Montreal Neurological Institute standard brain template. The middle column shows the association between the mean activation change of each cluster and Principal Component 3 'phonology' score change, using robust regression. For clusters B, C and E, the right column shows the association between the mean activation change of each cluster and Principal Component 3 'phonology' score change, controlling for baseline 'phonology' score, using robust regression. For clusters A, D and F, the right column shows the interaction plot for the adjusted association between the mean activation change of each cluster and Principal Component 3 'phonology' score change, with baseline 'phonology' score fixed at high, middle and low values, using robust regression. Statistical thresholding used a voxel-wise cluster forming threshold of $p<0.005$ (uncorrected) and a cluster-level threshold of $p<0.05$ after familywise error correction. Abbreviations: B = unstandardised regression coefficient; PC = Principal Component; PC3 = 'phonology' Principal Component; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post-stroke).

Discussion

This study investigated the nature of aphasia recovery post-stroke by analysing data from one of the largest longitudinal, combined neuropsychological and functional neuroimaging studies in the subacute stage of aphasia recovery to date. The results show for the first time that (i) behavioural variation is multidimensional at the subacute phase, (ii) recovery is also multidimensional in nature and does not reflect a single, unitary factor, and (iii) these multidimensional shifts in performance are related to changes in task-induced activation in different brain regions. Moreover, from a methodological perspective, the study also showed that investigations of recovery need to account for baseline performance to avoid potential false positive and negative findings. Overall, this work provides novel insights into the neural basis of aphasia recovery, and suggests that neuromodulatory treatments targeting the same neural region in all aphasic stroke survivors might be ineffective or even impair recovery, and should instead be tailored depending on the specific language profile of each individual patient.

First, considering the cross-sectional neuropsychological data from the subacute phase, we found that the PSA language variations can be dissociated into orthogonal principal components. Aligning with previous research (Kummerer et al., 2013), the observed PCA structure consisted of fluency (PC1), semantic/executive function (PC2) and phonology (PC3). These behavioural factors also directly mirror the core PCA structure obtained in previous investigations of PSA patients in the chronic phase (Butler et al., 2014; Halai, Woollams, et al., 2020; Halai et al., 2017; Lambon Ralph et al., 2010; Mirman et al., 2015). This suggests that, even though patients' performance improves over time, the behavioural variation in both the subacute and chronic phases can be captured by the same graded dimensions (phonology, fluency, semantics and executive skill).

A central aim of this study was to investigate whether these distinct underlying components of language have different recovery trajectories. The clear and novel answer was that recovery is not a single monolithic factor, but rather that the distinct language components recovered in either uncorrelated (for fluency vs. semantic/executive and semantic/executive vs. phonology) or anticorrelated (for fluency vs. phonology) ways. This result suggests that aphasia recovery is heterogenous and multidimensional. Consequently, aphasia treatment strategies and assessment of their efficacy should avoid single targets and outcome measures, and instead consider multiple, distinct aspects of language. This alternative approach might improve treatment outcomes and make efficacy measures more sensitive. It would also ensure that potential negative trade-offs between recovery of different language components are not overlooked.

A third key finding was that these different underlying language components are associated with changing activation in multiple, primarily non-overlapping regions during aphasia recovery. These associations can be positive or negative depending on the language component in question. After controlling for baseline performance, demographic and clinical variables: fluency recovery was associated with increasing activation in bilateral MFG and right temporooccipital MTG/supramarginal gyrus; semantic/executive recovery was associated with reducing activation in bilateral ATLs; while phonology recovery was associated with reducing activation in precentral gyri, dorso-medial frontal poles, and the precuneus. Overlapping clusters in the ventromedial prefrontal cortex were positively associated with fluency recovery but negatively associated with semantic/executive and phonology recovery. Thus, different aspects of language might rely on different neural regions for recovery. Unfortunately, clinical trials of non-invasive brain stimulation frequently target the same neural region in all stroke survivors regardless of their aphasia profile (Heikkinen et al., 2019; Heiss et al., 2013; Seniow et al., 2013; Waldowski et al., 2012). This study suggests that doing so might be entirely ineffective or even impair recovery, depending on the specific language profile of each individual patient. Similarly, meta-analyses of non-invasive brain stimulation in post-stroke aphasia tend to ignore the aphasia profile or stimulation site of included studies (Li et al., 2020; Zhang et al., 2021). Unless these are considered, potential benefits or harms of non-invasive brain stimulation might be overlooked.

This study was also able to explore an important methodological issue: one should control for baseline language performance when identifying neuroimaging correlates of aphasia recovery. This follows from the fact that, when task performance/PCA factor scores have greater variance in the acute than more chronic phases (e.g., when performance approaches ceiling, cf. control performance levels), then the change in the score is inescapably anti-correlated with the score in

the acute phase (Hope et al., 2019). With this in mind, we found little evidence for 'false positives' in the present study; most identified fMRI clusters were still significantly associated with PC change after controlling for T1 PC score. However, we did find multiple examples of 'false negatives', in which a true association between activation change and PC change was masked by varying baseline language performance. Strikingly, no regions were associated with semantic/executive change before controlling for T1 score, yet bilateral ATL clusters were negatively associated with semantic/executive recovery after controlling for T1 score.

An unexpected implication of this study is that activation changes during the subacute phase of aphasia recovery may not be as large as previously suggested, at least for patients with mild-moderate aphasia (Cardebat et al., 2003; Heiss et al., 1999; Long et al., 2018; Mattioli et al., 2014; Nenert et al., 2018; Saur et al., 2006; Stockert et al., 2020; van Oers et al., 2018). Specifically, we found no significant group-level activation differences between T1 and T2 post-stroke, despite significant behavioural improvement in 14 of the 16 neuropsychological tests. It is possible that this absence of group-level activation changes might be another consequence of the multidimensional nature of aphasia recovery. If recovery in different aspects of language relies on changing activation in different neural regions, then the resultant heterogeneity in activation changes will be much less likely to generate a single, homogeneous change when the data are combined simply at the group level.

Previous studies have reported regions of hyperactivation in PSA relative to controls at subacute timepoints post ictus (Mattioli et al., 2014; Nenert et al., 2018; Saur et al., 2006). We only found that, averaged across T1 and T2, patients hypoactivated the left posterior cingulate/precuneus, right thalamus and right temporo-parietal cortex. Less activation in the right inferior parietal part of this hypoactivated cluster was associated with poorer fluency at T1, while normalisation of activation was associated with better fluency recovery. This suggest that 'functional diaschisis' (Carrera & Tononi, 2014) in right inferior parietal cortex might contribute to fluency deficits, and their resolution, post-stroke. Although inferior parietal and posterior temporal regions are not classically associated with fluency, one model postulates that an efference copy of speech is sent from premotor to inferior parietal and posterior temporal regions for comparison with perceived auditory information and optimisation of ongoing speech (Rauschecker & Scott, 2009). Indeed, the right posterior superior temporal sulcus plays a role during speech production in controls (Yamamoto et al., 2019); changing grey matter integrity in the right posterior MTG has been associated with changing object naming in chronic aphasia (Hope et al., 2017); and grey matter

volume in the right temporoparietal cortex has been associated with spontaneous speech, naming and repetition scores in chronic aphasia (Xing et al., 2016).

Our results have wider implications for the mechanisms that might underlie aphasia recovery.

First, increasing and decreasing activation in specific regions of both hemispheres can be associated with recovery of different aspects of language. This is inconsistent with simplified left vs. right 'regional hierarchy' models wherein engagement of the right hemisphere is suboptimal to left-dominant activation or might even be maladaptive through 'transcallosal disinhibition' (Heiss & Thiel, 2006; Thiel et al., 2013). Instead, it is consistent with a model in which healthy language relies on a bilateral albeit asymmetric left-biased network that can support at least partial recovery through upregulation of function in perilesional and contralateral areas (for a recent computational implementation see (Chang & Lambon Ralph, 2020).

Secondly, activation change in the ventromedial prefrontal cortex, dorso-medial frontal poles and precuneus was negatively correlated with phonological recovery. These clusters overlapped with regions deactivated in controls; thus, greater deactivation in task-negative regions was associated with better phonological recovery. This is consistent with previous work demonstrating that greater differential activity between the default mode network and task-positive networks was associated with better language performance post-stroke (Geranmayeh et al., 2016).

Thirdly, activation change in bilateral MFG was positively associated with fluency recovery, yet these regions were not activated in controls performing the same task. One would need controls to perform a more difficult language task to know whether these regions represent spare capacity that is upregulated through variable neuro-displacement both pre-morbidly (when task difficulty increases) and after stroke (Geranmayeh et al., 2014; Stefaniak et al., 2020). A network including the bilateral MFG was activated during the 'Decision' task in controls (Geranmayeh et al., 2014), suggesting their involvement in domain-general executive processing (Geranmayeh et al., 2014). It is unclear why increased activity in domain-general regions was associated with fluency recovery but not semantic/executive recovery. A possible explanation is that the 'Speech+Count>Rest' contrast used in this analysis favours activations related to fluency, and the degree of recruitment of domain-general regions to a fluency task might not necessarily relate to the degree of recruitment of domain-general regions to tasks favouring semantic/executive function. Future work using separate fluency, semantic/executive and phonological fMRI tasks would be needed to investigate this further.

Finally, activation change in the bilateral ATLs was negatively associated with semantic/executive recovery. These regions have been proposed to comprise the 'hub' of a

distributed network subserving semantic representation (Lambon Ralph et al., 2017). Previous research has suggested that variable neuro-displacement enables intrinsic spare capacity within the bilateral ATLs to ameliorate semantic impairment following damage/stimulation to one part of the system (Binney & Lambon Ralph, 2015; Jung & Lambon Ralph, 2016; Rice et al., 2018). The negative association between activation change and semantic/executive recovery might reflect the fact that individuals with deteriorating efficiency of the distributed semantic network (and thus worse semantic/executive recovery) need to activate their 'hub' more to perform a given semantic task (Barulli & Stern, 2013; Nyberg et al., 2014), just as ATL activations in healthy participants tend to be higher for more demanding concepts or semantic tasks (Rice et al., 2018).

This study has several limitations. First, the specific neural regions that were associated with recovery on different language components in this study should be replicated in an independent sample and assessed regarding how well they predict recovery in individuals with aphasia. Second, analyses were restricted to voxels in which no patient had a lesion, in order to remove any direct confounding effect of variable lesion morphology. Consequently, we might have missed positive associations between language performance and left hemisphere activation, as has been shown previously (Stockert et al., 2020). However, most previous studies reporting such a compensatory role for left hemisphere language activation have tended to overlook variable lesion morphology as a confounding factor (Fridriksson, 2010; Saur et al., 2006; Szaflarski et al., 2013; C. A. M. M. van Oers et al., 2010), or analysed restricted subgroups of patients with lesions confined to certain locations (Heiss et al., 1999; Stockert et al., 2020) which does not completely accounting for voxelwise variability in lesion location throughout the entire left hemisphere. Third, previous work using independent component analysis identified that propositional and non-propositional speech can have opposite effects on activation in the same spatiotemporal networks in the left versus right hemispheres (Geranmayeh et al., 2014). Although we observed bilateral fronto-temporo-parietal activation and identified neural correlates for all three language components using 'Speech+Count>Rest', our use of this contrast of interest and the mass univariate analysis method does mean that we are unable to say which aspects of propositional or non-propositional speech, and which distinct spatiotemporal networks, might be contributing to any of the identified neural correlates of language change or to the observed negative associations between activation and PC2/PC3.

Conclusions and future directions

Our findings demonstrate that distinct underlying components of language have different recovery trajectories associated with changing activation in distinct neural regions. Targeting the

same neural region in all aphasic stroke survivors might be ineffective or even impair recovery, depending on the specific language profile of each individual patient. As noted above, given the significant clinical challenges that arise immediately after stroke, most studies including the current one tend to recruit patients with mild-moderate aphasia. Future studies are needed to explore the recovery profiles and their neural correlates in more severely affected patients.

Chapter 6

General Discussion

Since this thesis is presented in the journal format, in which Chapters 2-5 are written in the style of journal papers, there has already been discussion of chapter-specific issues within each chapter. This final Chapter therefore consists of several sections. First, the results of core Chapters 2-5 are drawn together in relation to the overarching themes discussed in Chapter 1. The two key neural compensatory mechanistic frameworks highlighted in Chapter 2, degeneracy and variable neurodisplacement, are then discussed before an additional theme repeatedly highlighted by the results of this thesis, 'diaschisis', is addressed. This Chapter concludes with a discussion of theoretical implications, clinical implications, suggested future research directions and limitations.

Overarching research themes

Theme 1 - a coherent mechanistic framework of aphasia recovery

The overarching purpose of this thesis was to better understand the neural compensatory mechanisms that take place to minimise language dysfunction, and promote language recovery, in post-stroke aphasia. The first theme called for a coherent mechanistic framework within which the multitude of existing recovery hypotheses can be conceptualised and from which specific predictions can be made that are refutable through empirical observation. Chapter 2 addresses the first research theme by proposing that existing theories of aphasia recovery can be conceptualised as specific examples of two more fundamental principles, degeneracy and variable neurodisplacement (Stefaniak et al., 2020). Degeneracy (Price & Friston, 2002) proposes that networks, which are not involved in the language task premorbidly, might become engaged after damage, either immediately or following experience-dependent plasticity (Stefaniak et al., 2020). Variable neurodisplacement (Manring & Johnson, 1996) proposes that spare capacity within or between neural networks might be downregulated in health under standard demand levels to save neural resources, but upregulated when demand increases (such as in post-stroke aphasia or difficult language tasks in health) (Stefaniak et al., 2020). These two mechanistic frameworks could apply across different aspects of not only aphasia recovery but, theoretically, any aspect of higher cognition. Importantly, degeneracy and variable neurodisplacement might occur irrespective of the specific anatomical regions or networks that might become utilised to aid language performance post-stroke and as such are mechanistic frameworks rather than verbal anatomical descriptions (Stefaniak et al., 2020).

For instance, right hemisphere regions involved in language post-stroke might be degenerate networks if they were language-specific homologues and quiescent premorbidly (Blank et al., 2003; Finger et al., 2003), or if they supported a non-language activity premorbidly but adapted to perform language computations through experience-dependent plasticity and 'neurocomputational invasion' post-stroke (Keidel et al., 2010). Alternatively, right hemisphere regions involved in language post-stroke might represent spare capacity if it were possible to observe their involvement during language pre-morbidly under situations of increased task difficulty (Stefaniak et al., 2020). Thus, both degeneracy and variable neurodisplacement might involve language-specific systems or non-language, domain-general cognitive systems involved in cognitive control and executive function (Brownsett et al., 2014; Stefaniak et al., 2020), including in the right hemisphere (Geranmayeh et al., 2014).

Importantly, we can derive specific predictions from this mechanistic framework that can be refuted through empirical observation. Both degeneracy and variable neurodisplacement predict that participants with aphasia should have regions of greater functional engagement during language than controls, and that the degree of engagement in such regions should be positively associated with language performance. Thus, under both mechanisms, we should be able to identify regions that compensate following damage to the language network. However, degeneracy would predict that such upregulated compensatory regions in post-stroke aphasia should not be engaged during language in health, even under increased task difficulty (Stefaniak et al., 2020). Conversely, variable neurodisplacement would predict that such upregulated compensatory regions should be engaged during language in health controls, particularly when task difficulty is increased (Stefaniak et al., 2020). As such, we might expect compensatory regions to more closely align with cognitive control networks under variable neurodisplacement than under degeneracy. However, most existing hypotheses, including the mechanistic framework outlined in Chapter 2, are still descriptive. There is, therefore, an ongoing need to develop computational models explaining how these mechanisms might be implemented, and for concrete experimental evidence to adjudicate between which mechanistic frameworks might occur *in vivo*.

Theme 2 - the neural basis of aphasia recovery

The second theme argues that we need to determine the neural regions that are functionally involved in language and its recovery post-stroke in order to provide empirical evidence for or against the proposed mechanisms of aphasia recovery. Functional involvement during

language might be determined through mass univariate activation-based imaging, through information-based multivariate pattern analysis, and through longitudinal studies that identify how regions functionally involved in language change over time.

Chapter 3 addresses the second research theme by investigating which specific regions are more or less likely to be activated during language in post-stroke aphasia than controls through an Activation Likelihood Estimation meta-analysis of coordinate-based language functional neuroimaging data for 481 individuals with post-stroke aphasia and 530 linked controls (Stefaniak et al., 2021). In so doing, it helped to adjudicate between which of the above proposed recovery mechanisms occur *in vivo*. Chapter 3 found that multiple regions throughout both hemispheres were consistently activated during language in both post-stroke aphasia and controls (Stefaniak et al., 2021). Thus, right hemisphere activation during language post-stroke should not be assumed to represent the recruitment of previously quiescent regions, since the healthy pre-morbid language network is inherently bilateral. Chapter 3 also concluded that the right anterior insula, frontal operculum and inferior frontal gyrus were more likely to be activated across all language tasks in post-stroke aphasia than controls, and crucially, that similar regions were more likely to be activated during higher than lower demand comprehension or production tasks in post-stroke aphasia and controls (Stefaniak et al., 2021). Since such regions appear to be engageable in healthy controls and particularly when task difficulty increases, Chapter 3's findings are most consistent with enhanced utilisation of spare capacity within right hemisphere executive regions via variable neurodisplacement (Stefaniak et al., 2021). However, since we were not able to test whether activation likelihood of such regions correlated positively with language performance, a limitation of Chapter 3 is that we are not able to definitively state that such language network changes are compensatory as opposed to being merely epiphenomenal. Unexpectedly, Chapter 3 found that multiple undamaged regions were less likely to be activated during language in post-stroke aphasia than controls, including domain-general regions of medial superior frontal cortex, right inferior frontal gyrus pars triangularis, and right temporal pole (Stefaniak et al., 2021). These changes might represent functional diaschisis (Carrera & Tononi, 2014), and also demonstrate that there is not global, undifferentiated upregulation of all domain-general neural regions during language in post-stroke aphasia (Stefaniak et al., 2021).

Chapter 4 addresses the second research theme by improving our understanding of the neural basis of language recovery using information processing as a proxy for functional

involvement. Chapter 4 was motivated by the fact that previous studies have focused almost exclusively on univariate language activation differences between post-stroke aphasia and controls (Stefaniak et al., 2021), but information-based multivariate pattern analysis might be more sensitive (Haxby et al., 2001; Haynes et al., 2007) at identifying regions functionally involved in language, particularly for higher cognitive functions such as language that involve distributed, overlapping representations (Carota et al., 2017). Chapter 4 presents one of the first language multivariate pattern analysis functional magnetic resonance imaging experiments (Fischer-Baum et al., 2017; Lee et al., 2017; Li et al., 2021) in patients with chronic post-stroke aphasia (n=24) and healthy controls (n=30). The primary language task consisted of passively listening to sentences with high versus low semantic predictability endings; a second scanner task consisted of hard versus easy visuospatial pattern matching to identify 'domain-general cognitive difficulty' regions. Clusters of significant sentence cloze decoding were present in core regions of the bilateral semantic control network (Jackson, 2021) (including bilateral inferior frontal gyri and medial superior frontal gyri) but additionally included regions that were not activated (left supramarginal gyrus) or even deactivated (precuneus) during the univariate language contrast, and might therefore have been overlooked were activation considered in isolation. There were no regions of significantly different language activation or information processing between participants with aphasia and controls at the group level. Language information content was significantly more positively associated with language score throughout bilateral fronto-temporo-parietal cortex in participants with aphasia than controls. This included regions of the bilateral inferior frontal gyrus pars opercularis/triangularis/orbitalis, bilateral dorsolateral and medial superior frontal gyrus, and bilateral middle temporal gyrus. However, these regions tended to have decreased decoding accuracy in participants with aphasia than controls. Taken together, Chapter 4 helps to adjudicate between which of the above proposed recovery mechanisms occur *in vivo*. Chapter 4 demonstrates that right hemisphere language information processing in post-stroke aphasia should not be assumed to represent the recruitment of previously quiescent regions through degeneracy, since the language network in health was found to be bilateral, and there was no evidence for the existence of novel regions that were not involved in language in controls but became involved during language post-stroke. Chapter 4's results are inconsistent with both degeneracy and variable neurodisplacement mechanisms of aphasia recovery, because both such mechanisms predict that regions in which information content is positively associated with language performance should have greater decoding accuracy in participants with aphasia than controls (i.e. be

upregulated to compensate for damage to the language network), whereas Chapter 4 found that such regions have lower decoding accuracy in participants with aphasia. Rather, Chapter 4's results suggest a novel form of 'information diaschisis' in which having a stroke was associated with lower language information processing in distant undamaged nodes of the residual language network which in turn contributed to the language impairment. Univariate language activation in these regions was not associated with language performance, showing that pattern decoding might be useful as a novel biomarker and that information content is more important than activation during aphasia recovery. Indeed, Chapter 4 suggests that it is not sufficient for a region to activate during recovery; regions must also process relevant information in order to contribute to behavioural performance. This is in keeping with a recent neurocomputational model of spoken language production and its recovery following virtual lesioning, which predicted that multivariate information content should be more tightly yoked to language performance than univariate activation (Chang & Lambon Ralph, 2020). Furthermore, in Chapter 4, most regions in which language information content correlated with language performance were domain-general, although their univariate activation during the pattern matching task was not associated with language. Chapter 4 therefore suggests that it is specifically the maintenance of language information processing, rather than univariate language activation or non-language activation, in a bilateral set of predominantly domain-general regions that is important for the maintenance of language performance post-stroke.

Theme 3 - the multidimensional nature of aphasia recovery

The third research theme argues that language is multidimensional and that distinct language dimensions might rely on different recovery mechanisms or changing functional involvement in different underlying neural regions post-stroke. Chapter 5 addressed the third research theme by analysing longitudinal functional magnetic resonance imaging and neuropsychological data from 26 individuals with post-stroke aphasia at 2 weeks and 4 months post-stroke. Principal Component Analysis revealed that patient language profiles in the subacute phase post-stroke were multidimensional in that they could be represented by three orthogonal components representing fluency, semantic/executive function and phonology. Crucially, these three components did not recover in a homogenous manner, but rather had uncorrelated recovery trajectories. Chapter 5 therefore demonstrates that aphasia recovery is heterogenous and multidimensional. Furthermore, different language components were associated with changing activation in multiple, non-overlapping bilateral

brain regions during aphasia recovery. Specifically, fluency recovery was associated with increasing activation in bilateral middle frontal gyri and right temporooccipital middle temporal gyrus; semantic/executive recovery was associated with reducing activation in bilateral anterior temporal lobes; while phonology recovery was associated with reducing activation in bilateral precentral gyri, dorso-medial frontal poles and the precuneus. Overlapping clusters in the ventromedial prefrontal cortex were positively associated with fluency recovery but negatively associated with semantic/executive and phonology recovery. These results provide novel insights into the multidimensional nature of aphasia recovery by demonstrating that different language dimensions seem to rely on changing functional involvement in different underlying neural regions post-stroke.

In addition to demonstrating the multidimensional nature of aphasia recovery, Chapter 5 provides insights into the neural basis of language recovery and the above proposed recovery mechanisms. The above observation, that increasing or decreasing activation in specific regions of both hemispheres can be associated with recovery of different aspects of language, is inconsistent with simplified left vs. right 'regional hierarchy' models in which engagement of the right hemisphere is suboptimal to left-dominant activation or even maladaptive through 'transcallosal inhibition' (Heiss & Thiel, 2006; Thiel et al., 2013). Indeed, Chapter 5 found that overt speech production resulted in significant bilateral activation in controls, which reiterates how right hemisphere language involvement in post-stroke aphasia should not be assumed to represent the recruitment of previously quiescent regions through degeneracy (Stefaniak et al., 2020) and is more consistent with a model in which healthy language relies on a bilateral albeit asymmetric left-biased network that can support partial recovery through upregulation of perilesional and contralateral areas (Chang & Lambon Ralph, 2020). Mean regional activation during overt speech production between 2 weeks and 4 months post-stroke was not significantly different between patient and control groups, suggesting that activation changes during the subacute phase of aphasia recovery may not be as large as previously suggested (Cardebat et al., 2003; Heiss et al., 1999; Stockert et al., 2020) and that future studies might explore using more sensitive multivariate information-based processing, as per Chapter 4, to identify regions of changing functional involvement in future longitudinal studies. Furthermore, although previous studies have reported regions of hyperactivation in post-stroke aphasia relative to controls at subacute timepoints (Mattioli et al., 2014; Nenert et al., 2018; Saur et al., 2006), Chapter 5 found no regions of significantly greater activation during language in participants with aphasia than controls. Since there was

no evidence, at the group level, for the existence of regions of greater activation during language in participants with aphasia than controls, Chapter 5's results do not provide evidence for degeneracy or variable neurodisplacement mechanisms of aphasia recovery, both of which predict that participants with aphasia should have regions of greater functional involvement during language than controls and that such regions should compensate for damage to the language network. Instead, Chapter 5 found that, averaged across 2 weeks and 4 months post-stroke, patients hypoactivated the left posterior cingulate/precuneus, right thalamus and right temporo-parietal cortex during overt speech production relative to controls. Less activation in the right inferior parietal part of this hypoactivated cluster was associated with poorer fluency at 2 weeks, while normalisation of activation was associated with better fluency recovery by 4 months, post-stroke. Chapter 5 therefore provides yet more evidence for the existence of 'functional diaschisis' (Carrera & Tononi, 2014) and its contribution towards language impairment in post-stroke aphasia.

Theoretical implications - neural compensatory mechanisms

Degeneracy and Variable Neurodisplacement

Both degeneracy and variable neurodisplacement predict that participants with aphasia should have regions of greater functional engagement during language than controls, and that the degree of engagement in such regions should be positively associated with language performance (i.e. compensate following damage to the language network). Degeneracy would predict that such 'upregulated' regions in post-stroke aphasia should not be engaged during language in health, even under increased task difficulty (Stefaniak et al., 2020). Conversely, variable neurodisplacement would predict that such regions should be engaged during language in health controls, particularly when task difficulty is increased (Stefaniak et al., 2020).

This thesis found surprisingly little evidence that participants with aphasia have regions of greater functional engagement during language than controls, which is a key prediction under both degeneracy and variable neurodisplacement mechanisms. Chapter 4 found no regions of significantly different language activation or information processing between participants with aphasia and controls at the group level; even in regions in which language information content was positively associated with language performance, decoding accuracy tended to be decreased (rather than increased) in participants with aphasia relative to controls. Chapter

5 also found no regions of significantly greater activation during overt speech production in participants with aphasia than controls when analysed at the group level. Only in the large-scale meta-analysis of Chapter 3 did we identify that the right anterior insula, frontal operculum and inferior frontal gyrus were more likely to be activated across all language tasks in post-stroke aphasia than controls (Stefaniak et al., 2021). Crucially, parts of the same right hemisphere regions were consistently activated across all language tasks in controls (Stefaniak et al., 2021) (suggesting that these upregulated regions are actually part of the pre-morbid language network), and similar regions were more likely to be activated during higher than lower demand comprehension or production tasks in participants with aphasia and controls (Stefaniak et al., 2021). Since such regions appear to be engageable in healthy controls and particularly when task difficulty increases, Chapter 3's findings are most consistent with them representing enhanced utilisation of spare capacity within right hemisphere executive regions via variable neurodisplacement (Stefaniak et al., 2021). However, since we were not able to test whether activation likelihood of such regions correlated positively with language performance, we are not able to definitively state that such language network changes are caused by compensatory mechanisms as opposed to being merely epiphenomenal.

Importantly, this thesis provides no evidence to suggest that degeneracy might cause compensatory language network changes in post-stroke aphasia. Chapters 3-5 each found that multiple regions throughout both hemispheres are functionally involved during language tasks in healthy individuals; this thesis therefore demonstrates that right hemisphere involvement during language post-stroke should not be assumed to represent the recruitment of previously quiescent regions, since the healthy pre-morbid language network is inherently bilateral. Thus, this thesis provides no evidence for the existence of novel regions being recruited into the language network de novo post-stroke and contributing towards language performance, consistent with the view that degeneracy mechanisms are not predominant during aphasia recovery.

Diaschisis

Unexpectedly, a recurring implication from all Chapters of this thesis is that diaschisis might play a role in post-stroke aphasia and its recovery. Diaschisis is the putative phenomenon, first termed by von Monakow in 1914, in which a focal brain lesion causes neurophysiological changes in distant, undamaged nodes of the affected cognitive network and this distant neurophysiological disruption in turns contributes to the degree of

impairment of cognitive performance (Carrera & Tononi, 2014). While reduced activation of distant undamaged regions has been described following focal subcortical (Baron et al., 1992) and cortical damage (Price et al., 2001), the degree to which the reduced activation correlates with behavioural performance (as opposed to being epiphenomenal) is unclear (Carrera & Tononi, 2014). More recently, a form of diaschisis termed 'connectional diaschisis' (Carrera & Tononi, 2014) has been described in which focal lesions are thought to result in connectivity changes between distant undamaged brain regions that in turn predict behavioural recovery (Grefkes & Fink, 2011; He et al., 2007; Siegel et al., 2016), although it is controversial as to whether knowledge of such connectivity changes adds anything to predictive models of behavioural performance beyond what would be explained by lesion location on its own (Salvalaggio et al., 2020) and thus whether such connectivity changes are truly relevant *in vivo* (Umarova & Thomalla, 2020). However, diaschisis has never hitherto been described in terms of changes to information processing in distant regions.

The meta-analysis in Chapter 3 found that multiple undamaged regions are less likely to be activated during language in post-stroke aphasia than controls, including domain-general regions of medial superior frontal and paracingulate cortex, right inferior frontal gyrus pars triangularis and right temporal pole (Stefaniak et al., 2021). Such changes are unlikely to be due to reduced 'cognitive effort' in participants with aphasia performing language tasks because other domain-general, demand-related regions of the right anterior insula and inferior frontal gyrus were more likely to be activated during language in participants with aphasia than controls (Stefaniak et al., 2021). These regions of reduced activation likelihood might therefore represent functional diaschisis, although since we were not able to test whether activation likelihood of such regions correlated positively with language performance, we are not able to state this with certainty (Stefaniak et al., 2021).

As stated previously, Chapter 5 provides evidence for the existence of functional diaschisis and its contribution towards language impairment in post-stroke aphasia. Averaged across 2 weeks and 4 months post-stroke, patients hypoactivated the right temporo-parietal cortex during overt speech production relative to controls. Less activation in the right inferior parietal part of this hypoactivated cluster was associated with poorer fluency at 2 weeks, while normalisation of activation was associated with better fluency recovery by 4 months, post-stroke. These changes are consistent with the notion that functional diaschisis, and its resolution, in the right inferior parietal cortex might contribute to dysfluency and its recovery in post-stroke aphasia. Although inferior parietal and posterior temporal regions are not

'classically' associated with fluency, the published literature does provide a theoretical and experimental basis to support its biologically plausibility (Hope et al., 2017; Rauschecker & Scott, 2009; Xing et al., 2016; Yamamoto et al., 2019).

However, it is Chapter 4 which provides the most compelling evidence for the existence of a novel form of diaschisis we term 'information diaschisis'. Chapter 4 found that language information content was significantly more positively associated with language performance throughout bilateral frontal and temporal regions (including inferior frontal gyrus pars opercularis/triangularis/orbitalis, dorsolateral and medial superior frontal gyrus, middle temporal gyrus) in participants with aphasia than controls. Crucially, these regions tended to have decreased decoding accuracy in participants with aphasia than controls, suggesting that rather than such regions being 'compensatory', we have instead identified a novel form of 'information diaschisis' in which having a stroke was associated with lower language information processing in distant undamaged regions which in turn contributed to the language impairment. This is the first time that diaschisis has been described in terms of changes to information processing in distant regions; as such, Chapter 4 adds significantly to our understanding of the mechanisms contributing to language impairment and its recovery post-stroke. Of note, diaschisis is a way through which aphasia becomes manifest post-stroke rather than a compensatory mechanism per se. However, it is possible that inter-individual differences in the propensity for diaschisis to develop and resolve might exist and contribute to inter-individual differences in patients' abilities to retain and recovery from aphasia post-stroke.

Theoretical implications - the neural basis of language

The bilateral language network

A consistent finding from all Chapters of this thesis is that multiple regions throughout both hemispheres are functionally involved during language tasks in healthy individuals, suggesting that the healthy pre-morbid language network is inherently bilateral. This strikes to the heart of an ongoing area of contention in the field of cognitive neuroscience, namely to what extent language computations are performed unilaterally or bilaterally (Stefaniak et al., 2020). Critics of a bilateral language network might argue that activation during a language task does not necessarily mean that behaviourally relevant neurocomputations are being performed and therefore that activation in right hemisphere regions during language might be epiphenomenal and superfluous to the language task itself. The contention that

regions 'necessary' for language should cause aphasia following damage has resulted in a historically strong 'left hemisphere only' view of language lateralisation, since the neuropsychological literature has tended towards a consensus that chronic aphasia results mainly from left hemisphere lesions (Geschwind, 1970; Kummerer et al., 2013). Some have even gone so far as to suggest that right hemisphere activation during a language task should be a considered a reason to question the validity of the language task being performed (Wilson et al., 2018).

However, a growing body of evidence, including this thesis, points towards a nuanced role for the right hemisphere during language in health and following aphasic stroke. First, it is now recognised that aphasia can result from isolated right hemisphere lesions (Gajardo-Vidal et al., 2018) or experimental deactivation (Hickok et al., 2008), demonstrating that the right hemisphere can be necessary for certain aspects of language function. Second, a recent neurocomputational model of spoken language production demonstrates that it is theoretically possible for the language network to be bilateral in healthy individuals yet mainly result in aphasia following left hemisphere damage, as a result of an intrinsically bilateral yet asymmetric language network (in which the computational capacity for language is greater in the left than in the right hemisphere) undergoing experience-dependent plasticity during recovery (Chang & Lambon Ralph, 2020). Thus, at least theoretically, it is possible for a region to contribute towards language function and be part of the 'language network' without being 'necessary' for normal language function (Chang & Lambon Ralph, 2020).

This thesis contributes to our understanding of the hemispheric distribution of the language network.

First, this thesis demonstrates, unequivocally, that right hemisphere regions are consistently activated during language in healthy individuals, at the group level. In the largest synthesis of the available functional neuroimaging literature in post-stroke aphasia and controls to date, Chapter 3 demonstrates that multiple regions throughout both hemispheres (including midline superior frontal gyrus, supplementary motor cortex and right frontal operculum, frontal orbital cortex, posterior superior temporal gyrus and posterior supramarginal gyrus) are consistently activated during language tasks in both post-stroke aphasia and controls (Stefaniak et al., 2021). Chapter 5 also found that overt speech production resulted in significant bilateral activation in controls.

Second, this thesis demonstrates that right hemisphere regions perform language information

processing in health individuals. Chapter 4 reported significant language information decoding in midline and right hemisphere regions (including right inferior frontal gyrus pars opercularis/triangularis/orbitalis, middle frontal gyrus, dorsolateral superior frontal gyrus and midline anterior cingulate cortex/supplementary motor area) and found no regions of significantly greater decoding in participants with aphasia than controls at the group level.

Third, this thesis associates changing language activation and information processing in various right hemisphere regions with changing language performance post-stroke. While not demonstrating causality, this suggests the involvement of at least some right hemisphere regions in the language deficit and its recovery following aphasic stroke.

Fourth, Chapter 4 found that diaschisis of language information processing in various right hemisphere regions correlates with the severity of the language deficit in participants with aphasia, but not healthy controls. This suggests that continued language information processing within right hemisphere parts of the healthy language network might be necessary for normal language performance, at least after damage to left hemisphere parts of the language network.

Involvement of domain-general versus domain-specific regions in language

Although the above discussion demonstrates the involvement of right hemisphere regions during language in health and post-stroke aphasia, it does not answer an ongoing point of contention in the literature regarding whether such regions are language-specific or domain-general. 'Language-specific' regions are putatively activated during language but not non-language tasks (Fedorenko et al., 2011; Pritchett et al., 2018). 'Domain-general' regions are a distributed set of cortical regions (including the bilateral dorsolateral prefrontal cortex, anterior insular cortex, pre-supplementary motor area, anterior/mid-cingulate gyrus, and intraparietal sulcus) (Fedorenko et al., 2013) that are activated during a wide variety of both language and non-language tasks that are cognitively demanding (Fedorenko et al., 2013) and as such have been termed the 'Multiple Demand' (MD) system (Duncan, 2010). MD regions are thought to mediate executive processes and are typically associated with performing explicit decision-making tasks that require behavioural control processes such as attention, strategy selection and performance monitoring (Fedorenko et al., 2013).

Right hemisphere activation during language has historically been assumed to represent the involvement of novel language-specific quiescent homologues via degeneracy (Blank et al.,

2003; Finger et al., 2003; Price & Friston, 2002). However, it has more recently been suggested that such right hemisphere regions activated during language in post-stroke aphasia represent regions involved in domain-general cognitive control (Geranmayeh et al., 2014). It is postulated that such regions become involved because more 'control' of goal-directed (language) behaviour is required when downstream language-specific regions are damaged (Geranmayeh et al., 2014). Supportive of this, evidence suggests that performance on behavioural tasks that assess top-down control frequently associates with language performance post-stroke (Murray, 2012; Rajtar-Zembaty et al., 2015; Su et al., 2015), while increased functional engagement of domain-general regions during language has been associated with longitudinal language recovery post-stroke (Geranmayeh et al., 2017). It has therefore been postulated that increased cognitive control, through upregulated involvement of domain-general regions, is a neural compensatory mechanism in post-stroke aphasia (Geranmayeh et al., 2014) with a particular emphasis placed on domain-general regions of the medial superior frontal cortex as a putative therapeutic target for non-invasive brain stimulation (Geranmayeh et al., 2017; Sliwinska et al., 2017). The main components of the existing narrative can therefore be summarised as follows: that domain-general executive regions are only minimally engaged during language in healthy participants (Fedorenko et al., 2011; Fedorenko et al., 2012) unless task difficulty is increased (Fedorenko et al., 2013; Stefaniak et al., 2020); that domain general regions are upregulated during language in post-stroke aphasia relative to healthy controls; that this is due to greater requirements for cognitive control following damage to domain-specific language regions and thus increased task difficulty; and that this serves to compensate following damage and support language performance in post-stroke aphasia (Geranmayeh et al., 2014; Geranmayeh et al., 2017).

This thesis supports the view that domain-general regions are functionally involved during language in participants with aphasia. For instance, Chapter 3 found overlap between the MD network (Fedorenko et al., 2013) and regions consistently activated by participants with aphasia during language in the left frontal lobe (frontal operculum, inferior frontal gyrus pars opercularis/triangularis, middle frontal gyrus), midline cortex (superior frontal gyrus, supplementary motor cortex) and right frontal lobe (frontal operculum, frontal orbital cortex) (Stefaniak et al., 2021). Furthermore, this thesis suggests that some domain-general regions might become more functionally engaged during language in post-stroke aphasia than controls. Chapter 3 found that there was overlap between the MD network (Fedorenko et al., 2013) and a region in the right anterior insula and frontal operculum which participants with

aphasia were more likely to activate during language than controls and during higher than lower demand language tasks (Stefaniak et al., 2021).

However, this thesis supports a more nuanced and, in some ways, novel role for domain-general regions during language in post-stroke aphasia than that contended by the existing narrative.

Firstly, rather than domain-general regions only being minimally engaged during language in healthy participants (Fedorenko et al., 2011; Fedorenko et al., 2012) unless task difficulty is increased (Fedorenko et al., 2013; Stefaniak et al., 2020), this thesis demonstrates that domain-general regions in both hemispheres and midline cortex are in fact consistently involved during language in healthy controls. Chapter 3 found overlap between the MD network (Fedorenko et al., 2013) and regions consistently activated by both participants with aphasia and healthy controls during language in the left frontal lobe (frontal operculum, inferior frontal gyrus pars opercularis/triangularis, middle frontal gyrus), midline cortex (superior frontal gyrus, supplementary motor cortex) and right frontal lobe (frontal operculum, frontal orbital cortex) (Stefaniak et al., 2021). Similarly, Chapter 5 found significant activation during overt speech production in healthy controls in regions including the bilateral middle frontal gyrus, inferior frontal gyrus pars opercularis, superior parietal lobule, and anterior cingulate cortex, which are all considered part of the MD network (Fedorenko et al., 2013).

Secondly, this thesis supports the novel and unexpected view that domain-general executive regions are functionally involved during language in healthy participants even during naturalistic language tasks with minimal explicit demands and thus no expectation that domain-general executive regions should be required. In Chapter 4, participants performed a task consisting of passively listening to sentences with high versus low semantic predictability endings. The sentence cloze decoding comparison did not require participants to behave differently or pay attention differently to high or low cloze sentences, and indeed, it was not explicitly revealed to participants that a contrast between these sentences was the purpose of the experiment. To ensure that participants were awake and attending to the auditory stimuli, participants were asked to press a button when the auditory stimuli told them to, but these infrequent button press trials were regressed out of the first level general linear model, and there was no additional task requiring participants to behave or respond differently to high versus low cloze sentences. As such, this task passively tapped into high-level aspects of semantic comprehension with no requirements for non-language executive

function, and this contrast should not, therefore, have required engagement of executive control. Nevertheless, Chapter 4 found that the bilateral inferior frontal gyrus pars opercularis/triangularis/orbitalis and medial superior frontal gyrus, regions which overlap with the MD network (Fedorenko et al., 2013), were activated when listening to sentences with lower more than higher semantically predictable endings. Chapter 4 also found that multivariate sentence cloze decoding was present in nine clusters and that six of these nine 'language information content' clusters, all located in midline or right hemisphere regions (midline supplementary motor area/medial superior frontal gyrus, right inferior frontal gyrus/opercularis/triangularis/orbitalis/middle frontal gyrus/dorsolateral and medial superior frontal gyrus), were significantly activated during a non-language executive task and thus were domain-general MD regions. Although it has previously been shown that MD regions become functionally engaged during more difficult than less difficult language task in healthy participants (Fedorenko et al., 2013), the observation that MD regions are functionally involved in a language task with no explicit behavioural task and minimal demands is novel and highly significant, and poses the question as to what function such domain-general regions serve during language tasks.

Thirdly, rather than all domain-general regions being upregulated in participants with aphasia relative to healthy controls (and this upregulation serving to compensate following damage and support language performance in post-stroke aphasia), this thesis demonstrates that it is instead more likely that most domain-general executive regions are less functionally engaged during language in participants with aphasia relative to healthy controls (and this functional disengagement correlates with, and might contribute to, the language impairment post-stroke via diaschisis). Chapter 3 found that domain-general regions of midline superior frontal gyrus and paracingulate gyrus, as well as (during production tasks) right frontal orbital cortex and anterior insula, both overlap with the MD network (Fedorenko et al., 2013) and are less likely to be activated during language in participants with aphasia than healthy controls (Stefaniak et al., 2021). Although Chapter 3 also found that different clusters in the right anterior insula, frontal operculum and inferior frontal gyrus pars opercularis were part of the MD network and more likely to be activated during language in participants with aphasia than healthy controls, the downregulation of other domain-general regions during language in post-stroke aphasia at least suggests that there is not global, undifferentiated upregulation of all domain-general regions during language and that there may be regional differences within the MD network in terms of recruitment and utility during language post-

stroke (Stefaniak et al., 2021). Crucially, Chapter 4 found that language information content (during a naturalistic sentence listening task designed so as to avoid requiring executive control) was reduced in post-stroke aphasia than controls in multiple regions (including the left middle frontal gyrus, dorsolateral superior frontal gyrus, putamen, cerebellum and bilateral inferior frontal gyrus opercularis/triangularis/orbitalis, precentral gyrus and insula); that the degree of reduction of language information content was associated with the degree of language impairment in participants with aphasia but not controls; and that these regions were significantly activated during a non-language cognitively demanding task (i.e. were domain-general control regions).

These results radically change our interpretation of the role that domain-general executive regions perform during language in health and post-stroke aphasia. They suggest that domain-general executive regions might perform a core function during language in healthy individuals that becomes necessary to maintain normal language performance in post-stroke aphasia. In keeping with this, a necessary role for domain-general executive regions during language in health was suggested by a previous lesion and functional imaging study which concluded that damage to the right inferior frontal sulcus results in comprehension impairment post-stroke because of this regions' involvement in working memory for both linguistic and non-linguistic material (i.e. domain-general executive function) in healthy individuals (Gajardo-Vidal et al., 2018). The current thesis extends this by suggesting that the role of domain-general executive regions during language does not relate solely to explicit decision making or typical 'executive' tasks, and that their involvement during language in post-stroke aphasia does not relate solely to increased requirements for cognitive control following damage to domain-specific language regions and increased task difficulty. Although MD regions are thought to mediate executive processes and are typically associated with performing explicit decision-making tasks that require behavioural control processes such as attention, strategy selection and performance monitoring (Fedorenko et al., 2013), Chapter 4 found that domain-general executive regions process high-level language information during a naturalistic sentence comprehension task with minimal executive demands in healthy individuals.

This poses the question as to what function such domain-general regions serve, and what is the nature of the computations they perform, during language tasks. One influential proposal suggests goal-directed behaviour is achieved by breaking down complex tasks into a series of sub-tasks that are focused on in turn (Duncan, 2010), and that MD cortex contributes to

this, in part, by holding a representation of the information that is currently being attended to which biases processing to attended or task-relevant information in other brain regions (Duncan, 2010). In this framework, a change to attended information should produce an update to the frontoparietal representation (Hon et al., 2006). Indeed, a previous study in healthy individuals instructed to watch a visual stream of words in a fixed location demonstrated that MD regions are activated when the attended sensory information (word) changed in the absence of any explicit task and thus no behaviour to perform (Hon et al., 2006). This mirrored earlier work showing that a right-lateralised network of regions (including the inferior frontal gyrus, insula, temporoparietal junction, cingulate and supplementary motor area) are activated during changes to the sensory environment across multiple modalities (Downar et al., 2000). Chapter 4 goes further by demonstrating that changes to the frontoparietal representation of currently attended information can be produced not only by low-level changes in simple stimuli but also by high-level linguistic changes in the semantic predictability of the last word of a sentence. It is possible that the ability of MD regions to process and/or detect high-level semantic changes to sentences is an important aspect of language processing that correlates with language performance following stroke.

Clinical implications and areas of future research

This research was performed because we hope that a better understanding of the neural compensatory mechanisms supporting the retention and recovery of language function post-stroke might enable the rational development of better therapeutic approaches for the management of patients with post-stroke aphasia. This thesis has clinical implications that should be accounted for when rationally developing therapeutic strategies that might use non-invasive brain stimulation to manipulate local neural function to facilitate language recovery post-stroke.

A conceptual implication from Chapter 4 is that it might be the local language information processing of a region, rather than its univariate activation, that is more important for language performance post stroke in multiple, bilateral neural regions. This is important when designing therapeutic strategies that attempt to manipulate neural activation to support language recovery through non-invasive brain stimulation technologies such as transcranial magnetic stimulation (Bucur & Papagno, 2019; Ren et al., 2014), because such technologies

are typically only thought of as being able to stimulate or inhibit the univariate activation of a region (Bestmann et al., 2005) or the distributed network or areas connected to the targeted region (Hodkinson et al., 2021; Jung et al., 2020). Encouragingly, recent work has shown that transcranial magnetic stimulation can alter locally multivariate information processing in addition to (Jackson et al., 2021), or even independently of (Rafiei et al., 2021), local univariate activation changes. However, the effects of non-invasive brain stimulation on information content decoding during language has not, to my knowledge, been investigated in the context of language in healthy individuals nor in post-stroke aphasia. Further research is needed to investigate this in the future.

A second conceptual implication from Chapters 3, 4 and 5 is that diaschisis likely contributes to language impairment in post-stroke aphasia and (from Chapter 5) that resolution of diaschisis might play a role in the recovery of language deficits subacutely. This suggests that previous attempts to aid language recovery by inhibiting right hemisphere neural regions such as the right inferior frontal gyrus pars triangularis (Ren et al., 2014; Thiel et al., 2013), supposedly so as to reduce transcallosal inhibition of dominant left-hemisphere language regions such as the left inferior frontal gyrus pars triangularis (Heiss & Thiel, 2006; Thiel et al., 2013), are conceptually flawed and that non-invasive brain stimulation strategies should be trialled that stimulate regions subacutely so to facilitate the recovery of diaschisis and, hopefully, language performance. Further research is needed to investigate whether stimulation of undamaged midline or right hemisphere regions during the sub-acute phase post-stroke is capable of facilitating the resolution of diaschisis and promote language recovery in future clinical trials.

Chapter 5 suggests that distinct underlying language components have different recovery trajectories associated with changing activation in distinct neural regions post-stroke. A conceptual implication from Chapter 5 is that there should not be a 'one size fits all' approach to neurorehabilitation using non-invasive brain stimulation, as targeting the same neural region in all aphasic stroke survivors might be ineffective or even impair recovery, depending on the specific language profile of each individual patient. Although not explicitly stated, this 'evidence-based medicine' approach to deciding whether non-invasive brain stimulation strategies work is undoubtedly being taken whenever a meta-analysis is published investigating whether 'TMS' or a specific frequency of TMS works (Li et al., 2020; Yao et al., 2020), without considering the stimulation site or patient profile of the participants included in each composite study and frequently considering 'TMS' to be a single,

homogenous treatment rather than being completely dependent on the specific neural region being targeted. We therefore need to move beyond 'group level' therapeutic strategies treating individuals with post-stroke aphasia as a homogenous group all receiving the same treatment. Chapter 5 calls for future clinical trials to take an individualised medicine approach in which each participant's specific language profile is accounted for when trialling, and ultimately deciding, the neural region to be targeted using non-invasive brain stimulation therapeutic strategies.

A final clinical implication from this thesis regards the neural site(s) that should be targeted using non-invasive brain stimulation to aid language recovery post-stroke. Ideally, there should be convergent evidence from multiple studies and multiple modalities for the potential involvement of a neural region during recovery for that region to be targeted using non-invasive brain stimulation. The limited number of participants in Chapters 4 and 5 reduces the certainty with which we can identify specific target sites from these studies in isolation. However, Chapter 3 is a meta-analysis of all available functional neuroimaging data in post-stroke aphasia to date and suggested that the right anterior insula, frontal operculum and inferior frontal gyrus are upregulated during language in post-stroke aphasia relative to controls while midline regions of the superior frontal gyrus are downregulated in participants with aphasia relative to controls via functional diaschisis (Stefaniak et al., 2021). Chapter 4 similarly found information diaschisis within midline regions of the superior frontal gyrus as well as, among other regions, the right inferior frontal gyrus. These studies would therefore raise the midline superior frontal gyrus and right inferior frontal gyrus as suitable neural regions that could be tested as therapeutic targets in future trials of non-invasive brain stimulation. However, Chapter 5 did not find positive associations between language recovery and the midline superior frontal gyrus nor the right inferior frontal gyrus. Indeed, Chapter 5 suggests that there shouldn't be a 'one size fits all' approach and that multiple target sites should be investigated for their effects on recovery of distinct underlying language dimensions. Ultimately, putative neural targets need to be tested in future work by robust, adequately blinded and adequately powered randomised controlled trials that take into account the language profile of each individual patient.

Limitations

Inevitably, the work within this thesis has limitations. Many have been mentioned within

each individual chapter's discussion sections, but some general, overarching limitations will be mentioned here.

A limitation of the mechanistic framework described in Chapter 2 is that we have not provided an implemented computational account of degeneracy or variable neurodisplacement and thus this framework is open to the accusation of being another verbal description of the mechanisms of aphasia recovery. A recent neurocomputational model (Chang & Lambon Ralph, 2020) provided an example of how variable neurodisplacement could be implemented using alternative language networks or pathways (Stefaniak et al., 2020), but a unified model of all potential types of degeneracy or variable neurodisplacement described in Chapter 2 has not been implemented. This highlights the need for further work to be performed in this area.

A limitation of Chapter 3 was that this meta-analysis was only able to look at activation differences but not correlate such findings with behavioural performance. It is thus possible that such language network changes are epiphenomenal and not subserving language performance or involved in language recovery. Neverthless, Chapter 3 adds significantly to our understanding of the language network changes that occur post-stroke and therefore is not without merit.

A related limitation is that this thesis used functional neuroimaging to identify regions functionally involved in language, rather than relying on lesion studies. A criticism is that we do not know whether activation is epiphenomenal, even if it correlates with language performance, and that lesion studies are needed to infer a necessary, causal role between the function of a region and behavioural performance. However, a recent neurocomputational model of spoken language production demonstrates that it is theoretically possible for the language network to be bilateral in healthy individuals yet mainly result in aphasia following left hemisphere damage (Chang & Lambon Ralph, 2020). Thus, at least theoretically, it is possible for a region to contribute towards language function and be part of the 'language network' without being 'necessary' for normal language function (Chang & Lambon Ralph, 2020). While lesion studies are important, it is therefore likely that they underestimate the extent of the neural regions involved in language post-stroke.

Finally, as mentioned above, the relatively limited number of participants in Chapters 4 and 5 reduces the certainty with which we can identify specific target sites from these studies in isolation. This highlights the difficulty recruiting individuals with post-stroke aphasia to functional neuroimaging studies, particularly in the subacute phase post-stroke, or to studies

with longitudinal designs. Further work is needed to recruit large cohorts of participants with aphasia to functional neuroimaging studies if this imaging modality is to be used as a prognostic tool in the future. Relatedly, group-level findings from Chapters 4 and 5 might not apply to individual participants with aphasia. However, these Chapters still add highlight important principles that advance our understanding of aphasia recovery post-stroke.

Appendices

Chapter 3 Supplementary Material

Supplementary Methods

Full search strategy and inclusion criteria

The databases Medline, Embase and PsycINFO were searched (using OvidSP for Medline and Embase) with the following criteria:

(aphasia OR dysphasia OR language OR fluency OR phonology OR semantics OR naming OR repetition OR comprehension OR speaking) AND (stroke OR ischaemia OR ischemia OR infarct) AND (fMRI OR PET OR neuroimaging OR imaging OR functional)

A search was performed with these criteria on 13/09/2017; we considered articles published at any time before this search date. This was updated on 17/03/2018 and again on 13/04/2020. After removal of duplicates, the title and abstracts of all search results were screened for potential relevance. Full-length manuscripts of all potentially relevant articles were reviewed. In addition, the bibliographies of all potentially relevant manuscripts were reviewed for additional articles that might have been missed in the primary search.

We identified eligible articles reporting observational studies that had: a) more than one person with language impairment at any time following left hemispheric stroke; b) more than one healthy control; and c) performed BOLD fMRI or ^{15}O-PET during a language task-based functional neuroimaging paradigm. Stroke survivors of both sexes, of any ethnicity, based in any clinical setting, with history of language impairment after haemorrhagic or ischaemic stroke were included. Studies with stroke survivors of any age were included as long as all aphasic group coordinates pertained to individuals who had stroke onsets when they were adults (>18 years old). Articles in which post-stroke aphasia was not specifically mentioned but in which stroke survivors had a documented language deficit at any time post-stroke were included. Studies in which stroke survivors initially had language impairments but whose language abilities had completely recovered were still included. We included studies regardless of blinding status. Conference abstracts, presentations, theses, and articles not written in English were excluded. We excluded studies without a comparator group or that related solely to alexia, agraphia, amusia, dysarthria, dysphonia, apraxia of speech or non-linguistic auditory processing but not aphasia. Articles using functional neuroimaging data from controls that was published in a separate paper were included as long as the control participant demographics and neuroimaging results were obtainable and the functional neuroimaging data from the aphasic group was novel. If the same stroke survivors with aphasia were included in functional

neuroimaging tasks in multiple papers, only the paper containing the largest number of participants was included. The only exceptions were: (1) if exactly the same aphasic and control participants were included in two papers and performed different functional neuroimaging tasks at the same time post-stroke that were analysed identically in the two papers, in which case the two papers were considered to be different functional neuroimaging tasks of the same group and were both included; or (2) if individual participant-level peak activation coordinates were available, in which case aphasia participants that were included in both papers were excluded from one of the papers.

We then extracted peak coordinate data from the list of articles that were eligible for inclusion in our meta-analysis. We extracted coordinate data for inclusion in this ALE meta-analysis that: related to activation (not deactivation) during a language task-based functional neuroimaging experiment; was provided in standard space; was derived from whole-brain mass-univariate analyses without region of interests (ROIs), small volume corrections (SVC), or conjunctions (Müller et al., 2018); and was calculated using the same significance thresholds in the aphasic group and control group. To be included, functional neuroimaging coordinates from the 'aphasic group' had to: exclude survivors of a right-sided stroke; exclude people without history of language impairment; and be collected before any study-specific research intervention might have occurred. Imaging analyses could not use masking, other than exclusive lesion masking or inclusive grey matter masking. In order to be as inclusive of the published literature as possible, we included coordinates obtained at any statistical significance threshold provided the same threshold was used in the aphasic group and control group of that article. If coordinates meeting these criteria for both the aphasic group and control group were not provided in the publication, authors were emailed for unpublished coordinates.

Coordinates from tasks at different timepoints on the same participant group were not pooled; only tasks performed at the longest timepoint post-stroke for each group were included. This was because we expect language network changes to occur between subacute and chronic stages in the same group, and we predicted that most functional neuroimaging studies would have been performed in the chronic phase post-stroke after language change had slowed or plateaued.

Subgroup ALE meta-analyses

PSA groups were divided into one of two categories according to a characteristic of its PSA participants or a characteristic of the imaging task performed. A contrast ALE meta-analysis compared the resultant two categories of PSA groups. Each control group was divided according to how its corresponding PSA group had been categorised and a contrast meta-

analysis performed on the resultant two categories of controls. If different PSA groups from within the same article possessed different patient characteristics and were divided into different categories, both groups were included in their respective categories and the paper's control coordinates were included in both categories for the control contrast meta-analysis. If the same PSA group performed multiple imaging tasks which were divided into different categories, the coordinates for both imaging tasks were included in their respective categories. Since contrast analyses were designed to look for regions of significantly different convergence between groups, the inclusion of coordinates from the same group in both subgroup datasets being contrasted would, if anything, reduce the likelihood of finding differences and thus should not increase the false positive rate.

Multiple Demand network overlay

An overlay of the Multiple Demand (MD) network was downloaded from 'http://imaging.mrc-cbu.cam.ac.uk/imaging/MDsystem' and corresponds to Figure 2 from Fedorenko et al. (2013) (Fedorenko et al., 2013). Group level t-statistics from seven different cognitively demanding tasks were used: left hemisphere data was reflected into the right; the resulting 14 t-maps were averaged and thresholded at t>1.5 (http://imaging.mrc-cbu.cam.ac.uk/imaging/MDsystem).

Semantic control network overlay

An overlay of the semantic control network, corresponding to Figure 1 from Jackson et al. (2021), was kindly provided by the article's author on personal request (Jackson, 2021). The semantic control network represents regions activated more during more controlled (harder) than less controlled (easier) semantic cognition in healthy individuals.

Statistical analysis

We compared the mean ages of the PSA and control groups using Mann-Whitney U tests implemented in SPSS version 25. Cohen's r^2 effect size and its associated 95% confidence interval were calculated using established methods (Cohen, 1988; Fritz et al., 2012).

Supplementary Results

Omnibus ALE meta-analysis across all language tasks

During all language tasks, PSA consistently activated bilateral regions, including: left frontal lobe (inferior frontal gyrus (IFG) pars opercularis/triangularis, frontal orbital cortex, middle frontal gyrus (MFG)); left temporal lobe (posterior middle temporal gyrus (MTG)); midline cortex (superior frontal gyrus (SFG), supplementary motor cortex (SMC), paracingulate gyrus); right frontal lobe (IFG pars opercularis/triangularis, frontal orbital cortex, precentral gyrus); right insula; and right temporal lobe (posterior superior temporal gyrus (STG), Heschl's gyrus, planum temporale) (Supplementary Table S5).

During all language tasks, controls consistently activated bilateral regions, including: left frontal lobe (IFG pars opercularis/triangularis, precentral gyrus, MFG); left insula; left temporal lobe (posterior MTG, posterior STG, planum temporale); left parietal lobe (superior parietal lobule, postcentral gyrus); midline cortex (SFG, SMC, paracingulate gyrus); right frontal lobe (frontal operculum, IFG pars triangularis, frontal orbital cortex); right temporal lobe (posterior STG, temporal pole); and right posterior supramarginal gyrus (Supplementary Table S6).

Comprehension tasks

During comprehension tasks, PSA consistently activated regions in: left frontal lobe (IFG pars opercularis/triangularis, frontal orbital cortex); left temporal lobe (posterior MTG); midline cortex (SFG, SMC); right frontal lobe (IFG pars triangularis, frontal orbital cortex, MFG); and right insula (Supplementary Table S9).

During comprehension tasks, controls consistently activated regions in: left frontal lobe (IFG pars opercularis/triangularis, MFG, frontal orbital cortex); left temporal lobe (temporooccipital ITG, posterior MTG); and midline cortex (SFG, paracingulate gyrus) (Supplementary Table S10).

Production tasks

During production tasks, PSA consistently activated regions in: left frontal lobe (IFG pars triangularis); midline cortex (SFG, SMC); right frontal lobe (IFG pars triangularis, precentral gyrus); right insula; and right temporal lobe (posterior STG, Heschl's gyrus) (Supplementary Table S13).

During production tasks, controls consistently activated regions in: left frontal lobe (IFG pars opercularis/triangularis, precentral gyrus, MFG); left insula; left temporal lobe (planum

temporale, temporooccipital MTG, posterior MTG); left insula; left posterior supramarginal gyrus; midline cortex (SFG, SMC, paracingulate gyrus); right posterior STG (Supplementary Table S14).

Comprehension>Production tasks

PSA were more likely to activate multiple regions during comprehension than production, including: left frontal lobe (IFG pars opercularis, frontal orbital cortex); left temporal lobe (posterior MTG); left parietal lobe (angular gyrus); right insula; and right frontal lobe (MFG) (Supplementary Table S18).

Production>Comprehension tasks

PSA were more likely to activate the following regions during production than comprehension: right frontal lobe (IFG pars opercularis, precentral gyrus); and right temporal lobe (posterior STG, planum temporale) (Supplementary Table S18).

Conjunction of PSA performing comprehension and production tasks

A conjunction demonstrated that PSA performing both comprehension and production tasks consistently activated overlapping regions in: left IFG pars triangularis; midline cortex (SFG, SMC); and right frontal lobe (frontal operculum cortex, frontal orbital cortex) (Supplementary Table S17).

Higher demand comprehension tasks

During higher demand comprehension tasks, PSA consistently activated regions in: left frontal lobe (IFG pars triangularis, frontal orbital cortex); left anterior STG; left superior lateral occipital cortex; midline cortex (SFG, SMC); right frontal lobe (IFG pars triangularis, frontal orbital cortex, MFG); and right insula (Supplementary Table S21).

During higher demand comprehension tasks, PSA and controls combined consistently activated regions in: left frontal lobe (frontal operculum, IFG pars opercularis/triangularis, frontal orbital cortex, precentral gyrus, MFG); left temporal lobe (temporooccipital ITG, posterior MTG); left insula; left occipital lobe (inferior lateral occipital cortex, superior lateral occipital cortex); midline cortex (SFG, SMC, paracingulate gyrus); right frontal lobe (IFG pars triangularis, frontal orbital cortex, MFG); right insula; and right occipital lobe (inferior lateral occipital cortex) (Supplementary Table S24).

Lower demand comprehension tasks

During lower demand comprehension tasks, PSA consistently activated regions in: left temporal lobe (temporal pole, posterior MTG) (Supplementary Table S22).

During lower demand comprehension tasks, PSA and controls combined consistently activated regions in: left frontal lobe (IFG pars triangularis, frontal orbital cortex); left temporal lobe (temporal pole, anterior MTG, posterior MTG); and right temporal pole (Supplementary Table S25).

Higher demand production tasks

During higher demand production tasks, PSA consistently activated regions in: left frontal lobe (IFG pars triangularis, MFG); right precentral gyrus; and right temporal lobe (Heschl's gyrus (includes H1 and H2), posterior STG) (Supplementary Table S27).

During higher demand production tasks, PSA and controls combined consistently activated regions in: left frontal lobe (frontal operculum, IFG pars opercularis/triangularis); left posterior MTG; midline cortex (SFG, SMC, paracingulate gyrus); right frontal lobe (frontal operculum, precentral gyrus); and right temporal lobe (Heschl's gyrus (includes H1 and H2), posterior STG) (Supplementary Table S30).

Lower demand production tasks

During lower demand production tasks, PSA consistently activated regions in: right precentral gyrus, right posterior STG (Supplementary Table S28).

During lower demand production tasks, PSA and controls combined consistently activated regions in: left temporal lobe (temporooccipital MTG, posterior STG); midline cortex (SFG, SMC, paracingulate gyrus); right precentral gyrus; and right posterior STG (Supplementary Table S31).

Search date	13/09/2017	17/03/2018	13/04/2020	Total
Number of articles after search of Medline and Embase	9221 (+9221)	746 (+746)	2095 (+2095)	12062 (+12062)
Number of articles after search of PsycINFO	11691 (+2470)	955 (+209)	2103 (+8)	14749 (+2687)
Number of articles after removal of duplicates	8015 (-3676)	363 (-592)	1791 (-312)	10169 (-4580)
Number of potentially relevant articles after screening title and abstract	317 (-7698)	21 (-342)	37 (-1754)	375 (-9794)
Number of potentially relevant articles after screening bibliographies	345 (+28)	21 (0)	40 (+3)	406 (+31)
Number of articles eligible for inclusion after assessment of full-length manuscript	67 (-278)	2 (-19)	10 (-30)	79 (-327)
Number of articles included in meta-analysis (appropriate coordinates obtained)	25 (-42)	2 (0)	6 (-4)	33 (-46)

Supplementary Table S1: Number of articles included and excluded at each stage of the systematic search. Numbers outside of parentheses represent the running total of the number of articles included at the corresponding row's stage of the systematic search, for the search date indicated by the corresponding column. Numbers in parentheses represent the number of articles added or removed from the systematic search between the previous and current row.

Search date	13/09/2017	17/03/2018	13/04/2020	Total
Not a full-length journal research article	102	7	6	115
No non-stroke control group	50	3	6	59
No functional neuroimaging in stroke survivors	35	3	7	45
Resting-state only	22	4	2	28
Significant neurological comorbidity	20	0	2	22
Includes stroke onset <18 years old	19	1	0	20
Data duplicated in another publication	14	1	3	18
No aphasia	8	0	1	9
Written in non-English language	6	0	1	7
Therapeutic intervention	1	0	2	3
Right-hemisphere stroke	1	0	0	1

Supplementary Table S2: Reasons for exclusion of potentially relevant articles after reading full length manuscript

Reference	Digital Object Identifier	Non-aphasic subgroup	Number in non-aphasic group	Non-aphasic group average age	PSA subgroup	Number in PSA group	PSA group average age	Months post stroke	Imaging task	Imaging modality	Task type	Task processing requirements
Allendorfer 2012	10.12659/MSM.882518	(only one non-aphasic group)	32	51.1	(only one PSA group)	16	54.4	44.4	Overt verb generation>Noun repetition	fMRI	Production	High
									Overt>Covert verb generation	fMRI	Production	High
Crinion 2005	10.1093/brain/awh659	(only one non-aphasic group)	18	58	Temporal patients	8	57.8	40.8	Listening to meaningful stories>Meaningless reversed speech	fMRI	Comprehension	Low
Crinion 2005	10.1093/brain/awh659	(only one non-aphasic group)	(18)	(58)	Aphasic control patients	9	66	48.1	(Listening to meaningful stories>Meaningless reversed speech)	(fMRI)	(Comprehension)	(Low)
Szaflarski 2011	10.1016/j.jstrokecerebrovasdis.2010.02.003	(only one non-aphasic group)	4	46	(only one PSA group)	4	53.3	57	Picture-name matching > Determining whether two geometric figures are identical	fMRI	Comprehension	Low
Mattioli 2014	10.1161/STROKEAHA.113.003192	(only one non-aphasic group)	10	(unknown)	(only one PSA group)	12	64.1	0.08	Intelligible sentence comprehension>Unintelligible sentence comprehension	fMRI	Comprehension	High
Qiu 2017	10.4103/1673-5374.198996	(only one non-aphasic group)	10	55.89	(only one PSA group)	10	55.9	1-3	Picture naming>Baseline	fMRI	Production	Low
Robson 2014	10.1093/brain/awt373	(only one non-aphasic group)	12	71	(only one PSA group)	12	70.1	20.3	Picture animate-inanimate judgement>Dual-baseline	fMRI	Comprehension	High
									Word animate-inanimate judgement>Dual-baseline	fMRI	Comprehension	High
Skipper-Kallal 2017	10.1155/2017/8740353	(only one non-aphasic group)	37	58.7	(only one PSA group)	39	59.8	52.9	Covert naming>Fixation	fMRI	Production	Low
									Overt naming>Fixation	fMRI	Production	Low
Blank 2003	10.1002/ana.10656	(only one non-aphasic group)	12	52.5	POp+	7	50	39	Propositional speech>Listening to environmental sounds	PET	Production	High
									Propositional speech>Counting aloud	PET	Production	High
Blank 2003	10.1002/ana.10656	(only one non-aphasic group)	(12)	(52.5)	POp-	7	61	17	(Propositional speech>Listening to environmental sounds)	(PET)	(Production)	(High)
									(Propositional speech>Counting aloud)	(PET)	(Production)	(High)

Cardebat 2003	10.1161/01.STR.0000099965.99393.83	(only one non-aphasic group)	6	50.6	(only one PSA group)	8	58.4	1.9 (first timepoint) 11.7 (last timepoint)	Word generation>Rest	PET	Production	High
Specht 2009	10.1016/j.neuroimage.2009.06.011	(only one non-aphasic group)	12	(unknown)	(only one PSA group)	12	49.8	22.9	Pseudoword decision task>Nonword decision task	PET	Comprehension	High
Weiller 1995	10.1002/ana.410370605	(only one non-aphasic group)	6	35	(only one PSA group)	6	57.7	43.2	Verb generation>Rest	PET	Production	High
									Verb generation>Pseudoword repetition	PET	Production	High
									Pseudoword repetition>Rest	PET	Production	High
									Pseudoword repetition>Verb generation	PET	Production	High
Fridriksson 2009	10.1002/hbm.20683	(only one non-aphasic group)	10	58.3	(only one PSA group)	11	58.8	37.6	Picture naming>Abstract colour picture viewing	fMRI	Production	Low
Warren 2009	10.1093/brain/awp270	(only one non-aphasic group)	11	54.6	(only one PSA group)	16	65.8	28.8	Listening to intelligible speech>Reversed speech	PET	Comprehension	Low
Sebastian 2012	10.1016/j.jneuroling.2012.01.003	Normal Control 1	1	57	Aphasic participant 1	1	60	76	Semantic judgement>Size judgement	fMRI	Comprehension	High
Sebastian 2012	10.1016/j.jneuroling.2012.01.003	Normal Control 2	1	(57)	Aphasic participant 2	1	53	12	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
Sebastian 2012	10.1016/j.jneuroling.2012.01.003	Normal Control 3	1	(57)	Aphasic participant 3	1	53	13	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
Sebastian 2011	10.1080/02687038.2011.557436	(only one non-aphasic group)	8	(unknown)	Aphasic participant 1	1	62	52	Semantic judgement>Size judgement	fMRI	Comprehension	High
									Picture naming>Scrambled picture viewing	fMRI	Production	Low
Sebastian 2011	10.1080/02687038.2011.557436	(only one non-aphasic group)	(8)	(unknown)	Aphasic participant 2	1	57	36	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
									(Picture naming>Scrambled picture viewing)	(fMRI)	(Production)	(Low)
Sebastian 2011	10.1080/02687038.2011.557436	(only one non-aphasic group)	(8)	(unknown)	Aphasic participant 3	1	60	78	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
									(Picture naming>Scrambled picture viewing)	(fMRI)	(Production)	(Low)
Sebastian 2011	10.1080/02687038.2011.557436	(only one non-aphasic group)	(8)	(unknown)	Aphasic participant 4	1	40	30	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
									(Picture naming>Scrambled picture viewing)	(fMRI)	(Production)	(Low)

Sebastian 2011	10.1080/02687038.2011.557436	(only one non-aphasic group)	(8)	(unknown)	Aphasic participant 5	1	70	36	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
									(Picture naming>Scrambled picture viewing)	(fMRI)	(Production)	(Low)
Sebastian 2011	10.1080/02687038.2011.557436	(only one non-aphasic group)	(8)	(unknown)	Aphasic participant 6	1	51	38	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
									(Picture naming>Scrambled picture viewing)	(fMRI)	(Production)	(Low)
Sebastian 2011	10.1080/02687038.2011.557436	(only one non-aphasic group)	(8)	(unknown)	Aphasic participant 7	1	79	56	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
									(Picture naming>Scrambled picture viewing)	(fMRI)	(Production)	(Low)
Sebastian 2011	10.1080/02687038.2011.557436	(only one non-aphasic group)	(8)	(unknown)	Aphasic participant 8	1	60	60	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
									(Picture naming>Scrambled picture viewing)	(fMRI)	(Production)	(Low)
Westmacott 2017	10.1017/cjn.2017.44	(only one non-aphasic group)	10	50.4	Adult stroke patient 1	1	45	24	Verb generation>Viewing symbol strings	fMRI	Production	Low
									Picture-word matching>Symbol matching	fMRI	Comprehension	Low
Westmacott 2017	10.1017/cjn.2017.44	(only one non-aphasic group)	(10)	(50.4)	Adult stroke patient 2	1	49	24	(Verb generation>Viewing symbol strings)	(fMRI)	(Production)	(Low)
									(Picture-word matching>Symbol matching)	(fMRI)	(Comprehension)	(Low)
Westmacott 2017	10.1017/cjn.2017.44	(only one non-aphasic group)	(10)	(50.4)	Adult stroke patient 3	1	57	12	(Verb generation>Viewing symbol strings)	(fMRI)	(Production)	(Low)
									(Picture-word matching>Symbol matching)	(fMRI)	(Comprehension)	(Low)
Westmacott 2017	10.1017/cjn.2017.44	(only one non-aphasic group)	(10)	(50.4)	Adult stroke patient 4	1	52	24	(Verb generation>Viewing symbol strings)	(fMRI)	(Production)	(Low)
									(Picture-word matching>Symbol matching)	(fMRI)	(Comprehension)	(Low)
Westmacott 2017	10.1017/cjn.2017.44	(only one non-aphasic group)	(10)	(50.4)	Adult stroke patient 5	1	43	24	(Verb generation>Viewing symbol strings)	(fMRI)	(Production)	(Low)
									(Picture-word matching>Symbol matching)	(fMRI)	(Comprehension)	(Low)
Abo 2004	10.1097/00001756-200408260-00011	(only one non-aphasic group)	6	21.33	Aphasic Case 1	1	56	60	Word repetition>Rest	fMRI	Production	Low
Abo 2004	10.1097/00001756-200408260-00011	(only one non-aphasic group)	(6)	(21.33)	Aphasic Case 2	1	55	60	(Word repetition>Rest)	(fMRI)	(Production)	(Low)
Perani 2003	10.1016/S0093-934X(02)00561-8		10	(unknown)	Aphasic case 1	1	46	12	Phonemic fluency>Rest	fMRI	Production	High

		(only one non-aphasic group)							Semantic fluency>Rest	fMRI	Production	High
Perani 2003	10.1016/S0093-934X(02)00561-8	(only one non-aphasic group)	(10)	(unknown)	Aphasic case 2	1	69	12	(Phonemic fluency>Rest)	(fMRI)	(Production)	(High)
									(Semantic fluency>Rest)	(fMRI)	(Production)	(High)
Perani 2003	10.1016/S0093-934X(02)00561-8	(only one non-aphasic group)	(10)	(unknown)	Aphasic case 3	1	65	10	(Phonemic fluency>Rest)	(fMRI)	(Production)	(High)
									(Semantic fluency>Rest)	(fMRI)	(Production)	(High)
Perani 2003	10.1016/S0093-934X(02)00561-8	(only one non-aphasic group)	(10)	(unknown)	Aphasic case 4	1	56	24	(Phonemic fluency>Rest)	(fMRI)	(Production)	(High)
									(Semantic fluency>Rest)	(fMRI)	(Production)	(High)
Rochon 2010	10.1016/j.bandl.2010.05.005	(only one non-aphasic group)	10	61	Untreated aphasic participant 1	1	83	30	Semantic judgement>Size judgement	fMRI	Comprehension	High
									Phonological judgement>Size judgement	fMRI	Production	High
Rochon 2010	10.1016/j.bandl.2010.05.005	(only one non-aphasic group)	(10)	(61)	Untreated aphasic participant 2	1	63	48	(Semantic judgement>Size judgement)	(fMRI)	(Comprehension)	(High)
									(Phonological judgement>Size judgement)	(fMRI)	(Production)	(High)
Sandberg 2014	10.1080/13554794.2013.770881	Normal healthy older adult 1	1	59.7	Aphasic participant 1	1	56	38	Word Judgement abstract words>Control items	fMRI	Comprehension	High
									Word Judgement concrete words>Control items	fMRI	Comprehension	High
									Word Judgement abstract words>Concrete words	fMRI	Comprehension	High
									Word Judgement concrete words>Abstract words	fMRI	Comprehension	High
									Synonym Judgement abstract words>Control items	fMRI	Comprehension	High
									Synonym Judgement concrete words>Control items	fMRI	Comprehension	High
									Synonym Judgement abstract words>Concrete words	fMRI	Comprehension	High
									Synonym Judgement concrete words>Abstract words	fMRI	Comprehension	High
Sandberg 2014	10.1080/13554794.2013.770881	Normal healthy older adult 2	1	(59.7)	Aphasic participant 2	1	55	76	(Word Judgement abstract words>Control items)	(fMRI)	(Comprehension)	(High)
									(Word Judgement concrete words>Control items)	(fMRI)	(Comprehension)	(High)
									(Word Judgement abstract words>Concrete words)	(fMRI)	(Comprehension)	(High)

									(Word Judgement concrete words>Abstract words)	(fMRI)	(Comprehension)	(High)
									(Synonym Judgement abstract words>Control items)	(fMRI)	(Comprehension)	(High)
									(Synonym Judgement concrete words>Control items)	(fMRI)	(Comprehension)	(High)
									(Synonym Judgement abstract words>Concrete words)	(fMRI)	(Comprehension)	(High)
									(Synonym Judgement concrete words>Abstract words)	(fMRI)	(Comprehension)	(High)
Sandberg 2014	10.1080/13554794.2013.770881	Normal healthy older adult 3	1	(59.7)	Aphasic participant 3	1	59	23	(Word Judgement abstract words>Control items)	(fMRI)	(Comprehension)	(High)
									(Word Judgement concrete words>Control items)	(fMRI)	(Comprehension)	(High)
									(Word Judgement abstract words>Concrete words)	(fMRI)	(Comprehension)	(High)
									(Word Judgement concrete words>Abstract words)	(fMRI)	(Comprehension)	(High)
									(Synonym Judgement abstract words>Control items)	(fMRI)	(Comprehension)	(High)
									(Synonym Judgement concrete words>Control items)	(fMRI)	(Comprehension)	(High)
									(Synonym Judgement abstract words>Concrete words)	(fMRI)	(Comprehension)	(High)
									(Synonym Judgement concrete words>Abstract words)	(fMRI)	(Comprehension)	(High)
Griffis 2017	10.1002/hbm.23476	(only one non-aphasic group)	43	54	(only one PSA group)	43	53	40.8	Semantic decision>Tone decision	fMRI	Comprehension	High
van Hees 2014	10.1016/j.bandl.2013.12.004	(only one non-aphasic group)	14	61.71	Aphasic participant 1	1	60	52.3	Picture naming>Passive viewing of scrambled line drawing	fMRI	Production	Low
van Hees 2014	10.1016/j.bandl.2013.12.004	(only one non-aphasic group)	(14)	(61.71)	Aphasic participant 2	1	60	(52.3)	(Picture naming>Passive viewing of scrambled line drawing)	(fMRI)	(Production)	(Low)
van Hees 2014	10.1016/j.bandl.2013.12.004	(only one non-aphasic group)	(14)	(61.71)	Aphasic participant 3	1	41	(52.3)	(Picture naming>Passive viewing of scrambled line drawing)	(fMRI)	(Production)	(Low)
van Hees 2014	10.1016/j.bandl.2013.12.004	(only one non-aphasic group)	(14)	(61.71)	Aphasic participant 4	1	52	(52.3)	(Picture naming>Passive viewing of scrambled line drawing)	(fMRI)	(Production)	(Low)

van Hees 2014	10.1016/j.bandl.2013.12.004	(only one non-aphasic group)	(14)	(61.71)	Aphasic participant 5	1	56	(52.3)	(Picture naming>Passive viewing of scrambled line drawing)	(fMRI)	(Production)	(Low)
van Hees 2014	10.1016/j.bandl.2013.12.004	(only one non-aphasic group)	(14)	(61.71)	Aphasic participant 6	1	48	(52.3)	(Picture naming>Passive viewing of scrambled line drawing)	(fMRI)	(Production)	(Low)
van Hees 2014	10.1016/j.bandl.2013.12.004	(only one non-aphasic group)	(14)	(61.71)	Aphasic participant 7	1	69	(52.3)	(Picture naming>Passive viewing of scrambled line drawing)	(fMRI)	(Production)	(Low)
van Hees 2014	10.1016/j.bandl.2013.12.004	(only one non-aphasic group)	(14)	(61.71)	Aphasic participant 8	1	65	(52.3)	(Picture naming>Passive viewing of scrambled line drawing)	(fMRI)	(Production)	(Low)
Schofield 2012	10.1523/JNEUROSCI.4670-11.2012	(only one non-aphasic group)	26	54.1	Moderate comprehension impairment	12	64.8	41.2	Listening to forward or reversed speech>Rest	fMRI	Comprehension	High
Schofield 2012	10.1523/JNEUROSCI.4670-11.2012	(only one non-aphasic group)	(26)	(54.1)	Severe comprehension impairment	9	57.0	37.4	(Listening to forward or reversed speech>Rest)	(fMRI)	(Comprehension)	(High)
Geranmayeh 2016	10.1212/WNL.0000000000002537	(only one non-aphasic group)	24	57	(only one PSA group)	53	62	3.7	Propositional speech production>Fixation	fMRI	Production	Low
									Propositional speech production>Counting	fMRI	Production	Low
									Propositional speech production>Go/no go button press	fMRI	Production	Low
Radman 2016	10.1155/2016/8797086	(only one non-aphasic group)	5	65.6	(only one PSA group)	4	63.4	1.0 (first timepoint)	Picture naming>Fixation	fMRI	Production	Low
								4.3 (last timepoint)				
Long 2017	10.1080/02687038.2017.1417538	(only one non-aphasic group)	5	56.6	(only one PSA group)	5	55.6	0.5-1.2 (first timepoint)	Passive reading>Passive checkerboard viewing	fMRI	Comprehension	Low
									Picture naming>Scrambled picture viewing	fMRI	Production	Low
Long 2017	10.1080/02687038.2017.1417538	(only one non-aphasic group)	(5)	(56.6)	(only one PSA group)	(5)	(55.6)	11-13 (last timepoint)	(Passive reading>Passive checkerboard viewing)	(fMRI)	(Comprehension)	(Low)
									(Picture naming>Scrambled picture viewing)	(fMRI)	(Production)	(Low)
Hallam 2018	10.1016/j.cortex.2017.10.004	(only one non-aphasic group)	16	64	(only one PSA group)	14	61	85.6	Passive listening to normal sentences>Spectrally rotated sentences	fMRI	Comprehension	Low

Nenert 2018	10.3233/RNN-170767	(only one non-aphasic group)	85	43	(only one PSA group)	14	46	0.5 (first timepoint)	Semantic decision>Tone decision	fMRI	Comprehension	High
									Covert verb generation>Finger tapping	fMRI	Production	High
Nenert 2018	10.3233/RNN-170767	(only one non-aphasic group)	(85)	(43)	(only one PSA group)	(15)	(46)	3 (last timepoint)	(Semantic decision>Tone decision)	(fMRI)	(Comprehension)	(High)
									(Covert verb generation>Finger tapping)	(fMRI)	(Production)	(High)
Stockert 2020	10.1093/brain/awaa023	(only one non-aphasic group)	17	51.9	Frontal lesions	17	52.3	0.1 (first timepoint)	Passive listening to normal sentences>Reversed sentences	fMRI	Comprehension	High
Stockert 2020	10.1093/brain/awaa023	(only one non-aphasic group)	(17)	(51.9)	(Frontal lesions)	(17)	(52.3)	9.1 (last timepoint)	(Passive listening to normal sentences>Reversed sentences)	(fMRI)	(Comprehension)	(High)
Stockert 2020	10.1093/brain/awaa023	(only one non-aphasic group)	(17)	(51.9)	Temporo-parietal lesions	17	54.4	0.1 (first timepoint)	(Passive listening to normal sentences>Reversed sentences)	(fMRI)	(Comprehension)	(High)
Stockert 2020	10.1093/brain/awaa023	(only one non-aphasic group)	(17)	(51.9)	(Temporo-parietal lesions)	(17)	(54.4)	8.8 (last timepoint)	(Passive listening to normal sentences>Reversed sentences)	(fMRI)	(Comprehension)	(High)
Meier 2019	10.1016/j.nicl.2019.101919	(only one non-aphasic group)	18	59.6	(only one PSA group)	34	61.9	60.0	Semantic feature judgement>Scrambled picture judgement	fMRI	Comprehension	High
Barbieri 2019	10.1016/j.cortex.2019.06.015	(only one non-aphasic group)	23	37.1	(only one PSA group)	16	48.1	49.1	Picture verification>Scrambled stimuli	fMRI	Comprehension	High
Tao 2019	10.1016/j.nicl.2019.101865	(only one non-aphasic group)	10	60.7	(only one PSA group)	15	61	58	Spelling task>Case verification task	fMRI	Comprehension	High
Wilson 2018	10.1002/hbm.24077	(only one non-aphasic group)	14	53.1	(only one PSA group)	15 (1 had bilateral strokes)	60.4	65.2	Narrative comprehension>Backwards speech	fMRI	Comprehension	Low
									Picture naming>Scrambled picture viewing	fMRI	Production	Low

Supplementary Table S3: Table of all PSA groups included in the ALE meta-analysis. Table of all included aphasic groups and their characteristics. Abbreviations: ALE = Activation Likelihood Estimation; fMRI = functional Magnetic Resonance Imaging; PET = Positron Emission Tomography; POp = pars opercularis; PSA = Post-stroke aphasia.

Supplementary Table S4: Table of all coordinates included in the omnibus ALE meta-analysis contrasting PSA versus controls performing all tasks. This table is located in a separate Excel file entitled 'Supplementary File 2', available through the figshare repository (doi:10.6084/m9.figshare.12582935). Table of all coordinates included in the omnibus ALE meta-analysis comparing all imaging tasks in all PSA vs. all imaging tasks in all controls. The paper, group, imaging contrast and number of participants are provided for all included coordinates at all timepoints at which functional neuroimaging was performed. Abbreviations: PSA = Post-stroke aphasia.

Single dataset meta-analysis:	1521 foci 64 participant groups 481 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	9224	-46	26	14	L Inferior Frontal Gyrus pars triangularis	45
		-48	14	16	L Inferior Frontal Gyrus pars opercularis	44
		-42	30	-8	L Frontal Orbital Cortex	47
		-42	10	30	L Middle Frontal Gyrus	9
2	7464	34	22	0	R Insular Cortex	*
		44	26	-8	R Frontal Orbital Cortex	47
		46	26	6	R Inferior Frontal Gyrus pars triangularis	13
		40	20	22	R Inferior Frontal Gyrus pars opercularis	9
		48	26	20	R Inferior Frontal Gyrus pars triangularis	46
		46	12	-4	R Insular Cortex	13
3	5568	-6	12	60	Superior Frontal Gyrus	6
		-4	2	60	Supplementary Motor Cortex	6
		-4	16	42	Paracingulate Gyrus	32
		10	16	44	Paracingulate Gyrus	32
		8	8	50	Supplementary Motor Cortex	6
4	4984	56	-30	4	R posterior Superior Temporal Gyrus	22
		64	-22	2	R posterior Superior Temporal Gyrus	41
		54	-18	10	R Heschl's Gyrus (includes H1 and H2)	41
		60	-8	0	R Planum Temporale	*
5	2376	-50	-38	0	L posterior Middle Temporal Gyrus	22
6	1616	54	-2	40	R Precentral Gyrus	6
7	1592	-46	-66	26	L superior Lateral Occipital Cortex	39

Supplementary Table S5: Significant clusters from the ALE map of all tasks in PSA. Table of clusters produced by the ALE single dataset meta-analysis of all imaging tasks in all PSA across all tasks. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; L=Left; MNI = Montreal Neurological Institute; R = Right.

Single dataset meta-analysis	809 foci 37 participant groups 530 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	12832	-48	12	12	L Inferior Frontal Gyrus pars opercularis	13
		-38	22	-2	L Insular Cortex	13
		-44	14	26	L Inferior Frontal Gyrus pars opercularis	9
		-46	30	12	L Inferior Frontal Gyrus pars triangularis	46
		-48	28	24	L Middle Frontal Gyrus	46
		-44	28	18	L Inferior Frontal Gyrus pars triangularis	46
		-56	-4	24	L Precentral Gyrus	6
		-50	-10	30	L Precentral Gyrus	6
		-46	30	-2	L Inferior Frontal Gyrus pars triangularis	13
2	5992	4	28	38	Paracingulate Gyrus	32
		-2	26	48	Superior Frontal Gyrus	8
		-2	28	40	Paracingulate Gyrus	32
		-2	12	56	Superior Frontal Gyrus	6
		-4	4	60	Supplementary Motor Cortex	6
		-6	18	48	Paracingulate Gyrus	6
		10	16	46	Paracingulate Gyrus	32
3	4080	-62	-42	4	L posterior Middle Temporal Gyrus	22
		-56	-42	4	L posterior Middle Temporal Gyrus	22
		-52	-34	8	L Planum Temporale	22
		-62	-30	0	L posterior Superior Temporal Gyrus	21
4	2816	38	24	-6	R Frontal Orbital Cortex	13
		58	12	-14	R Temporal Pole	22
		36	32	4	R Inferior Frontal Gyrus pars triangularis	45
		46	16	-4	R Frontal Operculum Cortex	13
5	1304	-46	4	52	L Middle Frontal Gyrus	6
6	1240	-34	-50	44	L Superior Parietal Lobule	40
		-40	-38	44	L Postcentral Gyrus	40
7	1168	64	-16	2	R posterior Superior Temporal Gyrus	22
		66	-28	4	R posterior Superior Temporal Gyrus	22
		66	-38	6	R posterior Supramarginal Gyrus	22

Supplementary Table S6: Significant clusters from the ALE map of all tasks in controls. Table of clusters produced by the ALE single dataset meta-analysis of all imaging tasks in all controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; L=Left; MNI = Montreal Neurological Institute; R = Right.

Conjunction

Cluster	Cluster size (mm^3)	x	y	z	Label	BA
1	6088	-46	30	12	L Inferior Frontal Gyrus pars triangularis	46
		-44	28	18	L Inferior Frontal Gyrus pars triangularis	46
		-48	14	16	L Inferior Frontal Gyrus pars opercularis	44
		-44	30	-4	L Frontal Orbital Cortex	47
		-36	28	-8	L Frontal Orbital Cortex	47
		-42	20	2	L Frontal Operculum Cortex	13
		-42	10	30	L Middle Frontal Gyrus	9
2	1544	-54	-40	2	L posterior Middle Temporal Gyrus	22
3	1176	38	24	-4	R Frontal Orbital Cortex	13
		48	18	-4	R Frontal Operculum Cortex	13
		38	26	2	R Frontal Operculum Cortex	13
4	824	64	-16	2	R posterior Superior Temporal Gyrus	22
		66	-26	2	R posterior Superior Temporal Gyrus	22
5	584	-4	14	56	Superior Frontal Gyrus	6
		-6	16	48	Paracingulate gyrus	32
6	464	-4	4	60	Supplementary Motor Cortex	6
7	304	-4	12	56	Superior Frontal Gyrus	6
8	200	10	16	46	Paracingulate gyrus	32
9	184	-4	8	60	Supplementary Motor Cortex	6
10	120	6	22	44	Paracingulate gyrus	6
		-2	22	42	Paracingulate gyrus	32
11	104	-6	4	60	Supplementary Motor Cortex	6
12	96	8	18	46	Paracingulate gyrus	6
13	88	-8	6	62	Supplementary Motor Cortex	6
14	72	-8	16	50	Superior Frontal Gyrus	6
15	64	-6	12	54	Superior Frontal Gyrus	6
16	64	-6	10	60	Superior Frontal Gyrus	6
17	48	-6	6	60	Supplementary Motor Cortex	6
18	40	-8	4	62	Supplementary Motor Cortex	6
19	32	6	20	44	Paracingulate Gyrus	6
20	32	-8	14	50	Paracingulate Gyrus	6
21	24	-4	-2	60	Supplementary Motor Cortex	6
22	16	6	22	42	Paracingulate Gyrus	6
23	16	8	14	44	Paracingulate Gyrus	32
24	8	62	-36	6	R posterior Supramarginal Gyrus	22
25	25	8	20	42	Paracingulate Gyrus	32
26	8	-8	12	50	Paracingulate Gyrus	6

Supplementary Table S7: Group similarities between the ALE maps of PSA and controls performing all tasks. Table of clusters produced by the omnibus ALE meta-analysis showing the conjunction between all imaging tasks in all PSA with all imaging tasks in all controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Abbreviations: ALE = Activation Likelihood Estimation.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
PSA>Controls							
1	280	40	20	12	2.32	R Inferior Frontal Gyrus pars opercularis	13
		40	18	6	2.05	R Frontal Operculum	13
Controls>PSA							
1	4024	-50	6	24	3.06	L Precentral Gyrus	6
		-52	12	10	2.99	L Inferior Frontal Gyrus pars opercularis	44
		-56	26	12	2.24	L Inferior Frontal Gyrus pars triangularis	45
		-58	-8	22	2.16	L Postcentral Gyrus	4
		-62	-4	22	2.15	L Precentral Gyrus	4
		-52	-12	30	1.91	L Postcentral Gyrus	6
		-42	14	32	1.70	L Middle Frontal Gyrus	9
2	2504	-2	28	46	2.70	Superior Frontal Gyrus	8
		-2	24	48	2.56	Superior Frontal Gyrus	8
		-4	30	34	2.37	Paracingulate Gyrus	32
		6	28	46	2.27	Superior Frontal Gyrus	8
		6	26	40	2.21	Paracingulate Gyrus	32
		-6	2	52	2.18	Supplementary Motor Cortex	6
		2	16	54	2.09	Superior Frontal Gyrus	6
		-2	8	54	2.04	Supplementary Motor Cortex	6
3	1968	-50	-34	12	2.82	L Planum Temporale	41
		-50	-30	10	2.67	L Planum Temporale	41
		-60	-48	4	2.64	L temporooccipital Middle Temporal Gyrus	22
4	1240	-34	-44	44	3.89	L Superior Parietal Lobule	40
		-30	-54	46	3.43	L Superior Parietal Lobule	7
		-40	-36	42	2.47	L Postcentral Gyrus	40
5	1152	-36	26	0	2.89	L Frontal Orbital Cortex	13
6	920	-44	6	56	2.97	L Middle Frontal Gyrus	6
7	592	58	14	-10	2.83	R Temporal Pole	22
8	528	-50	31	28	2.51	L Middle Frontal Gyrus	9
9	496	34	34	4	2.44	R Inferior Frontal Gyrus pars triangularis	45

Supplementary Table S8: Group differences between the ALE maps of PSA versus controls performing all tasks. Table of clusters produced by the omnibus ALE meta-analysis showing differences between all imaging tasks in all PSA vs. all imaging tasks in all controls. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Post-stroke aphasia.

Single dataset meta-analysis	1018 foci 40 participant groups 306 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	6216	-46	24	16	L Inferior Frontal Gyrus pars triangularis	46
		-46	14	16	L Inferior Frontal Gyrus pars opercularis	13
		-38	30	-10	L Frontal Orbital Cortex	47
2	3640	34	22	-2	R Insular Cortex	*
		44	26	-8	R Frontal Orbital Cortex	47
3	1896	-52	-38	0	L posterior Middle Temporal Gyrus	22
4	1664	-42	-64	26	L superior Lateral Occipital Cortex	39
5	1424	-6	16	52	Superior Frontal Gyrus	6
		-6	12	60	Superior Frontal Gyrus	6
		-2	0	60	Supplementary Motor Cortex	6
6	1160	40	22	22	R Middle Frontal Gyrus	9
		48	26	20	R Inferior Frontal Gyrus pars triangularis	46

Supplementary Table S9: Significant clusters from the ALE map of PSA performing comprehension tasks. Table of clusters produced by the ALE single dataset meta-analysis of comprehension tasks in PSA. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Post-stroke aphasia.

Single dataset meta-analysis	388 foci 25 participant groups 358 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	3968	-42	12	28	L Inferior Frontal Gyrus pars opercularis	9
		-48	14	14	L Inferior Frontal Gyrus pars opercularis	44
		-46	28	26	L Middle Frontal Gyrus	9
		-46	26	16	L Inferior Frontal Gyrus pars triangularis	46
		-44	24	30	L Middle Frontal Gyrus	9
2	1832	-46	32	-4	L Inferior Frontal Gyrus pars triangularis	47
		-38	30	-8	L Frontal Orbital Cortex	47
		-44	28	-20	L Frontal Orbital Cortex	47
3	1152	-4	48	34	Superior Frontal Gyrus	6
		-6	54	24	Superior Frontal Gyrus	9
		0	44	24	Paracingulate Gyrus	9
4	1104	-48	-52	-10	L temporooccipital Inferior Temporal Gyrus	37
		-46	-50	-16	L temporooccipital Inferior Temporal Gyrus	37
5	1088	-56	-40	2	L posterior Middle Temporal Gyrus	22
6	1024	-46	4	52	L Middle Frontal Gyrus	6

Supplementary Table S10: Significant clusters from the ALE map of controls performing comprehension tasks. Table of clusters produced by the ALE single dataset meta-analysis of comprehension tasks in controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right.

Conjunction						
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	2064	-48	14	16	L Inferior Frontal Gyrus pars opercularis	44
		-46	26	16	L Inferior Frontal Gyrus pars triangularis	46
		-46	28	22	L Inferior Frontal Gyrus pars triangularis	46
		-44	14	22	L Inferior Frontal Gyrus pars opercularis	9
2	720	-56	-42	2	L posterior Middle Temporal Gyrus	22
3	560	-36	28	-8	L Frontal Orbital Cortex	47
		-40	30	-6	L Frontal Orbital Cortex	47

Supplementary Table S11: Group similarities between the ALE maps of PSA and controls performing comprehension tasks. Table of clusters produced by the ALE contrast meta-analysis showing the conjunction between comprehension tasks in PSA with comprehension tasks in controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Abbreviations: ALE = Activation Likelihood Estimation; L=Left; MNI = Montreal Neurological Institute; R = Right.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
PSA>Controls (comprehension)							
1	424	34	18	4	2.20	R insular cortex	*
Controls>PSA (comprehension)							
1	712	-4	48	26	2.31	Paracingulate Gyrus	9
		-6	54	28	2.21	Superior Frontal Gyrus	9
		-4	54	20	2.06	Superior Frontal Gyrus	9
2	576	-44	4	52	2.25	L Middle Frontal Gyrus	6
3	352	-46	16	30	2.42	L Middle Frontal Gyrus	9
4	232	-52	32	-4	2.24	L Inferior Frontal Gyrus pars triangularis	47

Supplementary Table S12: Group differences between the ALE maps of PSA versus controls performing comprehension tasks. Table of clusters produced by the ALE contrast meta-analysis showing differences between comprehension tasks in PSA vs. comprehension tasks in controls. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Post-stroke aphasia.

Single dataset meta-analysis	503 foci 42 participant groups 225 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	5176	56	-30	4	R posterior Superior Temporal Gyrus	22
		64	-20	0	R posterior Superior Temporal Gyrus	*
		54	-18	10	R Heschl's Gyrus (includes H1 and H2)	41
		44	-22	12	R Heschl's Gyrus (includes H1 and H2)	41
2	2464	-46	32	14	L Inferior Frontal Gyrus pars triangularis	46
		-54	24	-2	L Inferior Frontal Gyrus pars triangularis	47
3	1920	50	-4	38	R Precentral Gyrus	6
		54	4	36	R Precentral Gyrus	6
4	1232	50	20	-4	R Inferior Frontal Gyrus pars triangularis	47
		46	12	-4	R Insular Cortex	13
5	1072	8	8	50	Supplementary Motor Cortex	6
		12	18	42	R Paracingulate Gyrus	32
6	808	-6	12	60	Superior Frontal Gyrus	6
		-4	-2	66	Supplementary Motor Cortex	6

Supplementary Table S13: Significant clusters from the ALE map of PSA performing production tasks. Table of clusters produced by the ALE single dataset meta-analysis of production tasks in PSA. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; L=Left; MNI = Montreal Neurological Institute; R = Right.

Single dataset meta-analysis	421 foci 18 participant groups 304 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	4952	-2	28	38	Paracingulate Gyrus	32
		-4	2	60	Supplementary Motor Cortex	6
		4	28	38	Paracingulate Gyrus	32
		10	16	46	Paracingulate Gyrus	32
		0	26	48	Superior Frontal Gyrus	8
		-4	12	52	Paracingulate Gyrus	6
		-4	2	70	Supplementary Motor Cortex	6
2	4680	-56	-4	24	L Precentral Gyrus	6
		-50	-10	30	L Precentral Gyrus	6
		-50	30	10	L Inferior Frontal Gyrus pars triangularis	46
		-50	12	10	L Inferior Frontal Gyrus pars opercularis	44
		-54	16	6	L Inferior Frontal Gyrus pars opercularis	44
		-52	20	8	L Inferior Frontal Gyrus pars opercularis	44
		-50	28	24	L Middle Frontal Gyrus	46
		-48	-12	40	L Precentral Gyrus	4
		-42	28	18	L Inferior Frontal Gyrus pars triangularis	46
		-48	8	20	L Inferior Frontal Gyrus pars opercularis	9
3	2560	-52	-34	8	L Planum Temporale	22
		-62	-42	6	L posterior Supramarginal Gyrus	22
		-56	-46	6	L temporooccipital Middle Temporal Gyrus	22
		-50	-40	-2	L posterior Middle Temporal Gyrus	22
4	1264	68	-28	2	R posterior Superior Temporal Gyrus	22
5	1080	-38	22	-2	L Insular Cortex	13

Supplementary Table S14: Significant clusters from the ALE map of controls performing production tasks. Table of clusters produced by the ALE single dataset meta-analysis of production tasks in controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; L=Left; MNI = Montreal Neurological Institute; R = Right.

Conjunction						
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	680	-48	30	12	L Inferior Frontal Gyrus pars triangularis	46
2	656	66	-22	2	R posterior Superior Temporal Gyrus	*
		62	-30	4	R posterior Superior Temporal Gyrus	22
3	376	10	16	44	Paracingulate Gyrus	32
4	312	-4	12	58	Superior Frontal Gyrus	6
		-4	8	60	Supplementary Motor Cortex	6
		-4	0	64	Supplementary Motor Cortex	6
5	8	66	-32	2	R posterior Superior Temporal Gyrus	22

Supplementary Table S15: Group similarities between the ALE maps of PSA and controls performing production tasks. Table of clusters produced by the ALE contrast meta-analysis showing the conjunction between production tasks in PSA with production tasks in controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Post-stroke aphasia.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
PSA>Controls (production)							
No clusters found							
Controls>PSA (production)							
1	6040	-54	0	22	3.16	L Precentral Gyrus	6
		-36	26	2	2.97	L Frontal Operculum Cortex	13
		-42	22	-4	2.75	L Frontal Orbital Cortex	13
		-52	10	12	2.73	L Inferior Frontal Gyrus pars opercularis	44
		-62	-4	22	2.64	L Precentral Gyrus	4
		-56	18	10	2.62	L Inferior Frontal Gyrus pars opercularis	44
		-54	26	12	2.54	L Inferior Frontal Gyrus pars triangularis	45
		-54	-12	30	2.44	L Postcentral Gyrus	4
		-50	8	22	2.37	L Precentral Gyrus	9
		-46	28	22	2.25	L Inferior Frontal Gyrus pars triangularis	46
2	4496	4	26	40	3.29	Paracingulate Gyrus	32
		6	18	54	2.89	Superior Frontal Gyrus	6
		2	22	50	2.70	Superior Frontal Gyrus	6
		-2	16	52	2.53	Superior Frontal Gyrus	6
		0	8	56	2.35	Supplementary Motor Cortex	6
		-2	4	62	2.29	Supplementary Motor Cortex	6
		-6	2	74	2.12	Superior Frontal Gyrus	6
		-6	2	54	2.06	Supplementary Motor Cortex	6
		14	16	50	1.81	R Superior Frontal Gyrus	6
3	2888	-48	-30	10	3.72	L Planum Temporale	41
		-56	-40	4	3.54	L posterior Superior Temporal Gyrus	22
4	952	68	-30	8	2.81	R posterior Superior Temporal Gyrus	42
5	560	36	26	-2	2.40	R Frontal Orbital Cortex	13
		36	28	-6	2.38	R Frontal Orbital Cortex	47
		38	20	-6	2.13	R Insular Cortex	47
6	488	48	-2	28	2.54	R Precentral Gyrus	6
		54	-4	26	2.37	R Precentral Gyrus	6
7	392	56	16	-10	2.64	R Temporal Pole	22
8	264	-68	-22	2	2.27	L posterior Superior Temporal Gyrus	22
9	200	52	-12	2	1.95	R Heschl's Gyrus (includes H1 and H2)	22

Supplementary Table S16: Group differences between the ALE maps of PSA versus controls performing production tasks. Table of clusters produced by the ALE contrast meta-analysis showing differences between production tasks in PSA vs. production tasks in controls. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Conjunction						
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	1008	-48	30	10	L Inferior Frontal Gyrus pars triangularis	46
2	448	-6	12	60	Superior Frontal Gyrus	6
		-4	2	62	Supplementary Motor Cortex	6
3	8	48	22	-6	R Frontal Orbital Cortex	47
4	8	46	22	-4	R Frontal Operculum Cortex	47
5	8	48	24	-4	R Frontal Orbital Cortex	47
6	8	-50	24	2	L Inferior Frontal Gyrus pars triangularis	45

Supplementary Table S17: Group similarities between the ALE maps of PSA performing comprehension tasks and production tasks. Table of clusters produced by the ALE contrast meta-analysis showing the conjunction between comprehension tasks in PSA with production tasks in PSA. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Post-stroke aphasia.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
Comprehension> Production (PSA)							
1	2128	-44	18	16	3.29	L Inferior Frontal Gyrus pars opercularis	46
2	1976	28	22	2	2.75	R Insular Cortex	*
		36	20	-8	2.73	R Insular Cortex	47
3	1080	-40	-68	20	3.09	L superior Lateral Occipital Cortex	39
		-36	-64	26	2.93	L Angular Gyrus	19
4	1040	-56	-39	-6	2.73	L posterior Middle Temporal Gyrus	*
5	1008	-32	28	-12	2.69	L Frontal Orbital Cortex	47
		-36	28	-14	2.66	L Frontal Orbital Cortex	47
6	384	44	28	24	2.13	R Middle Frontal Gyrus	9
Production> Comprehension (PSA)							
1	2216	60	-32	8	2.81	R posterior Superior Temporal Gyrus	41
		54	-26	8	2.49	R Planum Temporale	41
2	496	48	-10	36	2.43	R Precentral Gyrus	6
		52	-6	32	2.14	R Precentral Gyrus	6
3	224	51	16	0	2.05	R Inferior Frontal Gyrus pars opercularis	13

Supplementary Table S18: Group differences between the ALE maps of PSA performing comprehension tasks versus production tasks. Table of clusters produced by the ALE contrast meta-analysis showing differences between PSA performing comprehension tasks versus PSA performing production tasks. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Conjunction						
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	448	-60	-42	4	L posterior Middle Temporal Gyrus	22
		-56	-44	4	L temporooccipital Middle Temporal Gyrus	22
2	360	-50	12	12	L Inferior Frontal Gyrus pars opercularis	44
3	360	-46	28	22	L Inferior Frontal Gyrus pars triangularis	46
		-46	28	12	L Inferior Frontal Gyrus pars triangularis	46
		-44	28	18	L Inferior Frontal Gyrus pars triangularis	46
4	8	-50	20	12	L Inferior Frontal Gyrus pars opercularis	45
5	8	-52	20	14	L Inferior Frontal Gyrus pars opercularis	45

Supplementary Table S19: Group similarities between the ALE maps of controls performing comprehension tasks and production tasks. Table of clusters produced by the ALE contrast meta-analysis showing the conjunction between comprehension tasks in controls with production tasks in controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Post-stroke aphasia.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
Comprehension> Production (controls)							
1	1096	-44	32	-22	3.54	L Frontal Pole	47
		-44	32	-16	3.43	L Frontal Orbital Cortex	47
		-43	27	-22	3.35	L Temporal Pole	47
		-40	32	-18	3.12	L Frontal Orbital Cortex	47
		-46	34	-8	3.01	L Frontal Orbital Cortex	47
2	896	0	43	23	3.35	Paracingulate Gyrus	9
		0	44	28	3.29	Paracingulate Gyrus	9
		-2	44	36	2.65	Superior Frontal Gyrus	8
3	568	-48	-46	-18	2.40	L temporooccipital Inferior Temporal Gyrus	37
Production> Comprehension (controls)							
1	2096	-56	-6	26	3.54	L Precentral Gyrus	4
		-51	-6	29	3.16	L Precentral Gyrus	6
		-52	2	24	2.97	L Precentral Gyrus	9
		-48	-10	40	2.35	L Precentral Gyrus	4
2	1792	8	20	48	2.75	Paracingulate Gyrus	6
		6	18	52	2.67	Superior Frontal Gyrus	6
		0	24	40	2.21	Paracingulate Gyrus	32
3	1264	-48	-30	10	3.12	L Planum Temporale	41
		-52	-42	12	2.52	L posterior Supramarginal Gyrus	22
		-52	-48	6	1.89	L temporooccipital Middle Temporal Gyrus	21
4	1120	66	-27	0	3.89	R posterior Superior Temporal Gyrus	*
		68	-26	6	3.72	R posterior Superior Temporal Gyrus	22
5	1096	-2	1	67	3.09	Supplementary Motor Cortex	6
		-4	4	69	2.70	Supplementary Motor Cortex	6
6	440	-50	20	2	2.09	L Inferior Frontal Gyrus pars triangularis	44
		-56	18	10	1.88	L Inferior Frontal Gyrus pars opercularis	44
		-52	26	10	1.83	L Inferior Frontal Gyrus pars triangularis	45
7	336	-36	28	2	2.18	L Frontal Orbital Cortex	45
		-40	18	-4	1.83	L Insular Cortex	*
8	232	52	18	-8	2.51	R Temporal Pole	47

Supplementary Table S20: Group differences between the ALE maps of controls performing comprehension tasks versus production tasks. Table of clusters produced by the ALE contrast meta-analysis showing differences between controls performing comprehension tasks versus controls performing production tasks. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Single dataset meta-analysis	840 foci 28 participant groups 230 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	4904	34	22	-2	R Insular Cortex	*
		44	26	-8	R Frontal Orbital Cortex	47
2	4880	-46	24	16	L Inferior Frontal Gyrus pars triangularis	46
3	2056	-42	-64	26	L superior Lateral Occipital Cortex	39
4	1696	-6	16	52	Superior Frontal Gyrus	6
		-6	12	60	Superior Frontal Gyrus	6
		-2	0	60	Supplementary Motor Cortex	6
5	1576	40	22	22	R Middle Frontal Gyrus	9
		48	28	20	R Inferior Frontal Gyrus pars triangularis	46
6	904	-34	34	-14	L Frontal Orbital Cortex	47
		-42	30	-4	L Frontal Orbital Cortex	47
		-36	30	-10	L Frontal Orbital Cortex	47
7	816	-54	-10	-12	L anterior Superior Temporal Gyrus	22

Supplementary Table S21: Significant clusters from the ALE map of PSA performing higher demand comprehension tasks. Table of clusters produced by the ALE single dataset meta-analysis of higher demand comprehension tasks in PSA. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Single dataset meta-analysis	178 foci 12 participant groups 76 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	1632	-58	4	-20	L Temporal Pole	21
		-52	12	-20	L Temporal Pole	38
2	920	-52	-38	-2	L posterior Middle Temporal Gyrus	*
		-60	-40	-4	L posterior Middle Temporal Gyrus	21
3	880	-42	20	-32	L Temporal Pole	38

Supplementary Table S22: Significant clusters from the ALE map of PSA performing lower demand comprehension tasks. Table of clusters produced by the ALE single dataset meta-analysis of lower demand comprehension tasks in PSA. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
Higher Demand Comprehension>Lower Demand Comprehension (PSA)							
1	4904	40	24	2	3.24	R Frontal Operculum Cortex	13
		34	20	-6	2.85	R Insular Cortex	*
		30	18	2	2.67	R Insular Cortex	*
2	3064	-44	27	18	3.89	L Inferior Frontal Gyrus pars triangularis	46
		-41	22	18	3.72	L Inferior Frontal Gyrus pars opercularis	46
		-42	21	23	3.54	L Inferior Frontal Gyrus pars opercularis	46
3	1488	-42	-70	24	2.30	L superior Lateral Occipital Cortex	39
		-47	-70	22	2.25	L superior Lateral Occipital Cortex	39
		-42	-66	18	2.24	L superior Lateral Occipital Cortex	19
4	1376	45	27	18	2.69	R Inferior Frontal Gyrus pars triangularis	46
		40	22	18	2.61	R Inferior Frontal Gyrus pars triangularis	*
Lower Demand Comprehension> Higher Demand Comprehension (PSA)							
1	1584	-60	8	-26	2.45	L Temporal Pole	21
		-62	6	-22	2.45	L Temporal Pole	21
		-56	9	-23	2.42	L Temporal Pole	21
		-52	20	-20	1.88	L Temporal Pole	38
2	880	-44	17	-33	2.15	L Temporal Pole	38
		-43	23	-33	2.15	L Temporal Pole	38

Supplementary Table S23: Group differences between the ALE maps of PSA performing higher demand versus lower demand comprehension tasks. Table of clusters produced by the ALE contrast meta-analysis showing differences between PSA performing higher demand comprehension tasks versus PSA performing lower demand comprehension tasks. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Single dataset meta-analysis	1118 foci 46 participant groups 510 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	12776	-46	24	16	L Inferior Frontal Gyrus, pars triangularis	46
		-42	12	28	L Inferior Frontal Gyrus, pars opercularis	9
		-36	30	-8	L Frontal Orbital Cortex	47
		-42	20	2	L Frontal Operculum Cortex	13
		-30	22	2	L Insular Cortex	*
2	5864	34	22	-2	R Insular Cortex	*
		42	26	-8	R Frontal Orbital Cortex	47
3	5088	-4	10	60	Superior Frontal Gyrus	6
		-6	18	48	Paracingulate Gyrus	6
		-2	38	54	Superior Frontal Gyrus	6
		2	32	40	Paracingulate Gyrus	8
		4	28	40	Paracingulate Gyrus	32
		-2	2	50	Supplementary Motor Cortex	6
4	3608	-48	-52	-12	L temporooccipital Inferior Temporal Gyrus	37
		-52	-40	0	L posterior Middle Temporal Gyrus	22
		-48	-68	-12	L inferior Lateral Occipital Cortex	37
5	2752	-42	-64	26	L superior Lateral Occipital Cortex	39
		-46	-68	24	L superior Lateral Occipital Cortex	39
		-52	-66	34	L superior Lateral Occipital Cortex	39
		-40	-64	36	L superior Lateral Occipital Cortex	39
6	1680	-46	-2	52	L Precentral Gyrus	6
		-38	4	48	L Middle Frontal Gyrus	6
7	1208	40	22	24	R Middle Frontal Gyrus	9
		48	28	20	R Inferior Frontal Gyrus, pars triangularis	46
8	928	46	-82	-4	R inferior Lateral Occipital Cortex	19
		46	-74	-10	R inferior Lateral Occipital Cortex	19
		46	-72	-2	R inferior Lateral Occipital Cortex	19

Supplementary Table S24: Significant clusters from the ALE map of PSA and controls performing higher demand comprehension tasks. Table of clusters produced by the ALE single dataset meta-analysis of higher demand comprehension tasks in PSA and controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Single dataset meta-analysis	288 foci 19 participant groups 154 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	6280	-52	12	-18	L Temporal Pole	38
		-42	20	-34	L Temporal Pole	38
		-58	6	-20	L Temporal Pole	21
		-64	-8	-18	L anterior Middle Temporal Gyrus	21
		-62	-6	-22	L anterior Middle Temporal Gyrus	21
		-44	28	-20	L Frontal Orbital Cortex	47
		-46	10	-34	L Temporal Pole	38
2	3184	-54	-40	2	L posterior Middle Temporal Gyrus	22
3	2896	58	10	-38	R Temporal Pole	21
		50	16	-24	R Temporal Pole	38
		60	10	-18	R Temporal Pole	38
4	952	-58	32	0	L Inferior Frontal Gyrus, pars triangularis	45

Supplementary Table S25: Significant clusters from the ALE map of PSA and controls performing lower demand comprehension tasks. Table of clusters produced by the ALE single dataset meta-analysis of lower demand comprehension tasks in PSA and controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
Higher Demand Comprehension>Lower Demand Comprehension (PSA+Controls)							
1	5856	43	29	-4	3.35	R Frontal Orbital Cortex	47
		44	27	2	3.54	R Frontal Operculum Cortex	13
		39	23	1	3.43	R Frontal Operculum Cortex	13
		36	32	2	3.35	R Frontal Orbital Cortex	45
		34	21	-7	3.29	R Insular Cortex	*
		31	23	7	3.12	R Insular Cortex	*
		29	26	0	3.12	R Insular Cortex	*
2	5128	-39	19	23	3.89	L Inferior Frontal Gyrus, pars opercularis	9
		-39	18	16	3.72	L Inferior Frontal Gyrus, pars opercularis	13
		-36	30	16	3.12	L Inferior Frontal Gyrus, pars triangularis	*
		-44	16	6	3.06	L Inferior Frontal Gyrus, pars opercularis	44
		-46	36	10	3.01	L Inferior Frontal Gyrus, pars triangularis	46
		-52	18	5	2.99	L Inferior Frontal Gyrus, pars opercularis	44
		-40	18	4	2.83	L Frontal Operculum Cortex	13
3	3056	-6	8	54	2.17	Supplementary Motor Cortex	6
		2	34	44	2.15	Superior Frontal Gyrus	8
		-1	21	48	2.02	Paracingulate Gyrus	6
		-6	18	54	1.97	Superior Frontal Gyrus	6
		6	28	38	1.88	Paracingulate Gyrus	32
		-12	16	48	1.83	L Paracingulate Gyrus	32
4	1992	-42	-54	-10	2.52	L temporooccipital Inferior Temporal Gyrus	37
5	688	-34	4	44	2.76	L Middle Frontal Gyrus	6
		-34	0	46	2.72	L Middle Frontal Gyrus	6
		-42	-4	48	2.33	L Precentral Gyrus	6
6	616	36	22	22	2.22	R Inferior Frontal Gyrus, pars opercularis	9
7	400	42	-82	-2	2.06	R inferior Lateral Occipital Cortex	19
		46	-80	0	1.97	R inferior Lateral Occipital Cortex	19
8	256	-40	-70	22	2.26	L superior Lateral Occipital Cortex	39
Lower Demand Comprehension> Higher							

Demand Comprehension (PSA+Controls)							
1	5864	-58	12	-24	3.72	L Temporal Pole	38
		-54	6	-26	3.54	L Temporal Pole	21
		-45	14	-33	3.35	L Temporal Pole	38
		-44	24	-28	3.09	L Temporal Pole	47
		-45	22	-36	3.04	L Temporal Pole	38
		-64	-4	-20	2.69	L anterior Middle Temporal Gyrus	21
2	2720	60	8	-27	3.72	R Temporal Pole	21
		56	16	-28	3.54	R Temporal Pole	38
		58	14	-35	3.35	R Temporal Pole	38
		57	9	-36	3.29	R Temporal Pole	21
3	2096	-64	-44	5	3.06	L temporooccipital Middle Temporal Gyrus	22
4	568	-54	32	-2	2.09	L Inferior Frontal Gyrus, pars triangularis	45

Supplementary Table S26: Group differences between the ALE maps of PSA and controls performing higher demand comprehension tasks versus PSA and controls performing lower demand comprehension tasks. Table of clusters produced by the ALE contrast meta-analysis showing differences between PSA and controls performing higher demand comprehension tasks versus PSA and controls performing lower demand comprehension tasks. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Single dataset meta-analysis	178 foci 13 participant groups 118 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	1024	52	-18	10	R Heschl's Gyrus (includes H1 and H2)	41
		56	-26	4	R posterior Superior Temporal Gyrus	41
		44	-22	12	R Heschl's Gyrus (includes H1 and H2)	41
		50	-28	4	R posterior Superior Temporal Gyrus	41
		52	-24	2	R posterior Superior Temporal Gyrus	41
2	728	48	-6	36	R Precentral Gyrus	6
3	560	-48	32	10	L Inferior Frontal Gyrus pars triangularis	46
		-48	34	14	L Inferior Frontal Gyrus pars triangularis	46
		-54	34	20	L Middle Frontal Gyrus	46

Supplementary Table S27: Significant clusters from the ALE map of PSA performing higher demand production tasks. Table of clusters produced by the ALE single dataset meta-analysis of higher demand production tasks in PSA. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Single dataset meta-analysis	325 foci 29 participant groups 107 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	2896	56	-30	2	R posterior Superior Temporal Gyrus	22
		66	-22	-2	R posterior Superior Temporal Gyrus	21
2	1040	54	4	36	R Precentral Gyrus	6
		56	0	38	R Precentral Gyrus	6

Supplementary Table S28: Significant clusters from the ALE map of PSA performing lower demand production tasks. Table of clusters produced by the ALE single dataset meta-analysis of lower demand production tasks in PSA. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
Higher Demand Production> Lower Demand Production (PSA)							
1	416	48	-6	40	2.09	R Precentral Gyrus	6
2	328	50	-22	12	1.95	R Heschl's Gyrus (includes H1 and H2)	41
		56	-22	10	1.93	R Planum Temporale	41
3	288	52	16	0	2.16	R Inferior Frontal Gyrus pars opercularis	13
		48	18	2	2.08	R Frontal Operculum Cortex	44
Lower Demand Production> Higher Demand Production (PSA)							
No clusters found							

Supplementary Table S29: Group differences between the ALE maps of PSA performing higher demand versus lower demand production tasks. Table of clusters produced by the ALE contrast meta-analysis showing differences between PSA performing higher demand production tasks versus PSA performing lower demand production tasks. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Single dataset meta-analysis	367 foci 21 participant groups 303 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	4640	-50	30	10	L Inferior Frontal Gyrus, pars triangularis	46
		-52	20	6	L Inferior Frontal Gyrus, pars opercularis	44
		-48	14	8	L Inferior Frontal Gyrus, pars opercularis	44
		-52	24	-2	L Inferior Frontal Gyrus, pars triangularis	45
		-42	28	18	L Inferior Frontal Gyrus, pars triangularis	46
		-38	22	0	L Frontal Operculum Cortex	13
2	1864	50	-16	4	R Heschl's Gyrus (includes H1 and H2)	13
		58	-28	6	R posterior Superior Temporal Gyrus	22
3	1640	48	16	-4	R Frontal Operculum Cortex	13
4	1032	48	-6	38	R Precentral Gyrus	6
5	840	8	20	46	Paracingulate Gyrus	6
		0	26	48	Superior Frontal Gyrus	8
		8	10	50	Paracingulate Gyrus	6
6	824	-48	-40	0	L posterior Middle Temporal Gyrus	22
7	776	-2	6	64	Supplementary Motor Cortex	6

Supplementary Table S30: Significant clusters from the ALE map of PSA and controls performing higher demand production tasks. Table of clusters produced by the ALE single dataset meta-analysis of higher demand production tasks in PSA and controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Single dataset meta-analysis	557 foci 39 participant groups 226 participants					
Cluster	**Cluster size (mm^3)**	**x**	**y**	**z**	**Label**	**BA**
1	4256	56	-32	2	R posterior Superior Temporal Gyrus	22
		66	-24	2	R posterior Superior Temporal Gyrus	22
2	3304	-4	0	60	Supplementary Motor Cortex	6
		-4	-2	68	Supplementary Motor Cortex	6
		-4	12	54	Superior Frontal Gyrus	6
		-8	10	40	Paracingulate Gyrus	24
		-4	14	42	Paracingulate Gyrus	32
3	1968	-62	-44	6	L temporooccipital Middle Temporal Gyrus	22
		-52	-36	8	L posterior Superior Temporal Gyrus	22
4	1912	54	-4	40	R Precentral Gyrus	6

Supplementary Table S31: Significant clusters from the ALE map of PSA and controls performing lower demand production tasks. Table of clusters produced by the ALE single dataset meta-analysis of lower demand production tasks in PSA and controls. For each cluster, we provide peak MNI coordinates, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). The ALE map was thresholded with a voxel-wise uncorrected $p<0.001$ cluster-forming threshold and a cluster-wise family-wise error corrected threshold of $p<0.05$ based on 1000 random permutations. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Cluster	Cluster size (mm^3)	x	y	z	Z (peak)	Label	BA
Higher Demand Production> Lower Demand Production (PSA+Controls)							
1	3464	-55	18	2	3.29	L Inferior Frontal Gyrus, pars opercularis	44
		-46	16	4	3.12	L Inferior Frontal Gyrus, pars opercularis	44
		-44	12	10	2.82	L Inferior Frontal Gyrus, pars opercularis	13
		-54	26	12	2.26	L Inferior Frontal Gyrus, pars triangularis	45
		-50	34	6	2.01	L Inferior Frontal Gyrus, pars triangularis	46
		-54	34	14	1.90	L Inferior Frontal Gyrus, pars triangularis	46
2	1256	54	16	0	3.54	R Inferior Frontal Gyrus, pars opercularis	44
3	664	-44	-42	2	2.70	L posterior Middle Temporal Gyrus	19
4	568	48	-20	8	2.37	R Heschl's Gyrus (includes H1 and H2)	13
		50	-20	4	2.20	R Heschl's Gyrus (includes H1 and H2)	13
		54	-24	10	2.13	R Planum Temporale	41
5	344	42	30	0	1.86	R Frontal Orbital Cortex	45
Lower Demand Production> Higher Demand Production (PSA+Controls)							
1	240	-62	-50	6	2.50	L temporooccipital Middle Temporal Gyrus	22

Supplementary Table S32: Group differences between the ALE maps of PSA and controls performing higher demand production tasks versus PSA and controls performing lower demand production tasks. Table of clusters produced by the ALE contrast meta-analysis showing differences between PSA and controls performing higher demand production tasks versus PSA and controls performing lower demand production tasks. For each cluster, we provide the peak MNI coordinate, cluster size, anatomical label (defined according to the Harvard-Oxford atlas) and Brodmann Area (determined using the Talairach Daemon atlas). Thresholded ALE maps from the two datasets being compared were subtracted from each other and thresholded at $p<0.05$ using 10000 P-value permutations with a minimum cluster threshold of $200mm^3$. Abbreviations: ALE = Activation Likelihood Estimation; BA = Brodmann Area; L=Left; MNI = Montreal Neurological Institute; R = Right. PSA = Person(s) with Post-stroke aphasia. * = No Brodmann Area according to Talairach Daemon.

Dataset 1			Dataset 2					
	Number of participant groups (with mean age available, total)	**Mean ages of participant groups (median, IQR)**		**Number of participant groups (with mean age available, total)**	**Mean ages of participant groups (median, IQR)**	**Mann-Whitney U**	**Mann-Whitney Z**	**Cohen's r^2 (95% CI)**	**P value**
Controls, all language tasks	33 (37)	57.0 (8.2)	PSA, all language tasks	64 (64)	57.4 (9.0)	878.0	-1.36	0.02 (-0.03-0.07)	0.18
Controls, comprehension tasks	22 (25)	57.0 (6.9)	PSA, comprehension tasks	40 (40)	57.0 (9.5)	391.5	-0.71	0.008 (-0.04-0.05)	0.48
Controls, production tasks	16 (18)	54.5 (8.2)	PSA, production tasks	42 (42)	57.4 (10.3)	252.0	-1.46	0.04 (-0.06-0.13)	0.14
Controls, comprehension tasks	22 (25)	57.0 (6.9)	Controls, production tasks	16 (18)	54.5 (8.2)	142.0	-1.01	0.03 (-0.07-0.12)	0.33
PSA, comprehension tasks	40 (40)	57.0 (9.5)	PSA, production tasks	42 (42)	57.4 (10.3)	828.0	-0.11	0.0001 (-0.005-0.005)	0.91
PSA, higher demand comprehension tasks	28 (28)	58.0 (9.8)	PSA, lower demand comprehension tasks	12 (12)	56.3 (11.1)	141.5	-0.78	0.02 (-0.06-0.09)	0.44

PSA, higher demand production tasks	12 (12)	58.1 (13.4)	PSA, lower demand production tasks	30 (30)	57.0 (9.1)	160.0	-0.56	0.007 (-0.04-0.06)	0.59

Supplementary Table S33: Mean ages of the participant groups contributing to contrast ALE meta-analyses in this paper. This table compares the obtainable mean ages of participant groups in each pair of datasets for every contrast ALE meta-analysis in this paper. 'P value' corresponds to uncorrected two-sided p-values from Mann-Whitney U-tests comparing the mean ages of participant groups in Dataset 1 to the mean ages of participant groups in Dataset 2. Abbreviations: IQR = Interquartile Range.

Chapter 4 Supplementary Material

Supplementary Methods

Neuropsychological tests

The administered tests were: immediate word repetition (40 items from the PALPA 9) (Kay et al., 1992); immediate non-word repetition (30 items from the PALPA 8) (Kay et al., 1992); picture naming (including 32 items from the Cambridge Naming Test (Bozeat et al., 2000) and 30 items from the Boston Naming Test (BNT)) (Kaplan et al., 1983); word-to-picture matching (32 items from the Cambridge Semantic Battery (CSB)) (Bozeat et al., 2000); forward and backward digit span (Wechsler, 1987); synonym judgement (48 items from the Synonym Judgement Test) (Jefferies et al., 2009); spoken sentence comprehension (subtest of the Comprehensive Aphasia Test (CAT)) (Swinburn et al., 2005); tests of non-verbal executive function (Raven's Coloured Progressive Matrices (RCPM) (Raven, 1962) and Brixton Spatial Anticipation Test (BSAT)) (Burgess & Shallice, 1997); and measures extracted from the Boston Diagnostic Aphasia Examination (BDAE) 'Cookie Theft' picture description task (Goodglass & Kaplan, 1983) assessing connected speech quanta (number of speech tokens, words per minute, mean length of utterance) (details of coding as per previous publications) (Borovsky et al., 2007; Halai et al., 2017).

MRI scan

MRI data were acquired on a 3T Siemens Prisma scanner (Siemens Verio, Erlangen, Germany) with a 32-channel head coil. High resolution structural T_1-weighted images with whole-brain coverage were acquired using a 3D Magnetization-Prepared Rapid Gradient-Echo (MPRAGE) sequence (repetition time (TR)=2250ms, TE=3.02ms, inversion time 900ms, field of view 256x256x192mm, flip angle=9°, GRAPPA acceleration factor 2, slice thickness=1mm). Functional images were acquired using echoplanar imaging (EPI) with multiecho multiband acquisition (TR=1792ms, TE=13ms/25.85ms/38.7ms/51.55ms, voxel size 3x3x3mm, multiband factor 2).

Preprocessing

Raw DICOM data were converted to nifti format using dcm2niix (Li et al., 2016). T_1-weighted structural images were pre-processed using 'fsl_anat' in the FMRIB Software Library (FSLv5.0.11, www.fmrib.ox.ac.uk/fsl) (Jenkinson et al., 2002; Jenkinson et al., 2012) and

lesion segmented and normalised to MNI space using LINDA (v0.5.0) (Ito et al., 2019; Pustina et al., 2016). All images warped to MNI space were visually inspected for accuracy. The functional EPI data were preprocessed using a combination of tools in FSL, AFNI (v18.3.03) (Cox, 1996; Cox & Hyde, 1997) and a python package to perform multi-echo ICA denoising ('TE-dependent analysis', or 'tedana') (Kundu et al., 2013; Kundu et al., 2012). Despiked (3dDespike), slice time corrected (3dTshift) and realigned (3dvolreg) EPI images were submitted to the 'tedana' toolbox (v0.0.8), which optimally combined the signal across echoes (weighted according to the estimated T2* in each voxel) and denoised the data by decomposing it (with PCA then ICA, maximum iterations=500, maximum restarts=50), classifying the resulting components as either BOLD (signal) or non-BOLD (noise) based on their TE-dependence, and removing noise components from the data (Kundu et al., 2013; Kundu et al., 2012). The optimally-combined, denoised images were coregistered to the T1 (FLIRT) and normalised to MNI space (using transforms obtained from LINDA).

Voxel Based Correlational Methodology

VBCM analyses were restricted to left hemisphere grey matter voxels lesioned in at least one stroke patient by using a grey matter lesion overlap mask that was created with SPM's Imcalc as the overlap between the binarised group lesion overlap mask and the SPM grey matter tissue integrity prior thresholded at >0.5.

Functional MRI design

Participants with corrected-to-normal vision wore MRI-compatible goggles holding lenses of their equivalent prescription. A hearing check was performed at the beginning of each MRI session in which the EPI sequence ran for 60s (33 TRs) and example audio files were played; the volume was individually increased until the participant reported being able to hear all audio files comfortably over the scanner noise.

Listening task

The projection screen visible to each participant was a black background throughout. Scanning stopped in between each run. Participants were instructed to keep their eyes closed and press a button with their left hand when they were told to.

Sentences and pseudoword strings were created and recorded by the same male native English speaker (JDS) using Audacity software. All auditory stimuli were peak amplitude normalised to 0 dbFS.

Psycholinguistic properties of words were obtained using 'Neighbourhood Watch' software (Davis, 2005) based on the British National Corpus. Corresponding words in the high and low cloze sentences were matched for frequency (mean 22884 [SD 30438] per million in high vs. 22850 [SD 30457] per million in low cloze sentences; paired *t*-test, t_{198}=0.06, p_{uncorr}=0.96), subjective familiarity (mean 590 [SD 50] in high vs. 591 [SD 49] in low cloze sentences; paired *t*-test, t_{153}=-0.41, p_{uncorr}=0.68), number of syllables (mean 1.31 [SD 0.59] in high vs. 1.31 [SD 0.54] in low cloze sentences; paired *t*-test, t_{199}=0.0, p_{uncorr}=1.0), mean token biphone frequency (mean 22313 [SD 28626] in high vs. 22141 [SD 28734] in low cloze sentences; paired *t*-test, t_{196}=0.31, p_{uncorr}=0.76), phonological neighbourhood (mean 13.24 [SD 7.81] in high vs. 13.24 [SD 7.15] in low cloze sentences; paired *t*-test, t_{196}=0.0, p_{uncorr}=1.0), imageability (mean 378.21 [SD 172] in high vs. 383.71 [SD 173.66] in low cloze sentences; paired *t*-test, t_{153}=-1.26, p_{uncorr}=0.21), and their sentences were matched for duration (mean 1633ms [SD 137ms] in high vs. 1656ms [SD 131ms] in low cloze sentences; *t*-test, t_{78}=-0.76, p_{uncorr}=0.45).

To create pseudoword strings, each word used in the sentences was converted to a corresponding pseudoword using 'Wuggy' software (Keuleers & Brysbaert, 2010) and the order of pseudowords in each string was randomised. Pseudoword strings were spectrally rotated around a centre frequency of 2000Hz and noise vocoded in order to degrade phonological information in 'rotated-vocoded' stimuli.

Other tasks

Speech produced during the repetition task was recorded with an Optoacoustics FOMRI-III microphone. The repetition task consisted of a single run lasting 598s (334 TRs). Data from the repetition task was designed for use in an ongoing study and was not used in the present paper.

Functional MRI first-level analysis – mass univariate activation

SPM's FAST prewhitening option was used to account for temporal auocorrelation (Olszowy et al., 2019).

Functional MRI first-level analysis – multivariate information

There are several multivariate classifiers available (Misaki et al., 2010), most of which enable decoding either by identifying differences in mean univariate activation between high and low cloze sentences, or by identifying differences in the multivariate pattern of distributed activation between high and low cloze sentences.

It was not possible to perform MVPA on the ‘pattern-matching’ task because MVPA requires the collection of multiple independent runs to enable cross-validation with independent data, and the ‘pattern-matching’ task only involved a single run.

Functional MRI second-level analysis

Analyses were restricted to a grey matter mask (created using SPM’s Imcalc by thresholding SPM’s grey matter tissue integrity prior at >0.5).

Second-level analyses used the CANlab Robust Regression Toolbox (Wager et al., 2005) (https://github.com/canlab/RobustToolbox) and SPM12. This toolbox used robust regression with iteratively reweighted least squares in order to automatically downweight the influence of outliers at each voxel (Wager et al., 2005). Robust regression was used because it can have increased power and reduced type 1 error rates compared to standard OLS regression (Wager et al., 2005), and it can be broadly similar to permutation testing in terms of type 1 error rate but with greater power when outliers are present (Mumford, 2017). Correction for multiple comparisons used voxelwise FDR $q<0.05$ (two-sided) with a minimum of 20 significant voxels per cluster; voxelwise thresholding provided greater spatial specificity than cluster-level thresholding (Woo et al., 2014) while FDR did not require significant smoothing of the EPI.

Anatomical labels were defined using the Automated Anatomical Labelling atlas (Tzourio-Mazoyer et al., 2002).

Supplementary Results

Identifier	Age	Sex	Years education	Lesion volume (voxels)	Years post stroke
Post-stroke aphasia					
1	68	F	20	164,516	10
2	52	M	12	49,393	13
3	70	F	12	83,038	4
4	41	F	12	84,388	1
5	40	F	18	4,072	2
6	58	M	18	27,754	0.5
7	77	M	10	53,878	3
8	69	M	20	8,712	10
9	75	M	19	76,302	6
10	70	M	14	9,842	1
11	82	M	19	8,362	8
12	64	M	13	31,964	11
13	40	F	18	1,073	2
14	67	M	14	95,087	5
15	50	F	12	144,662	6
16	59	M	13	139,503	6
19	50	M	20	61,959	3.5
20	68	M	10	125,700	2
21	76	M	16	208,352	23
22	31	M	13	5,408	0.5
24	59	F	12	3,100	1
25	34	M	10	128,386	9
26	60	M	14	102,100	2
27	75	M	18	278,909	12
Controls					
1	53	M	17		
2	55	M	14		
3	64	F	20		
4	62	M	11		
5	69	F	15		
6	62	F	20		
7	58	M	17		
8	62	M	22		
9	64	M	16		
10	64	F	17		
11	64	F	12		
12	69	F	14		
13	75	M	13		

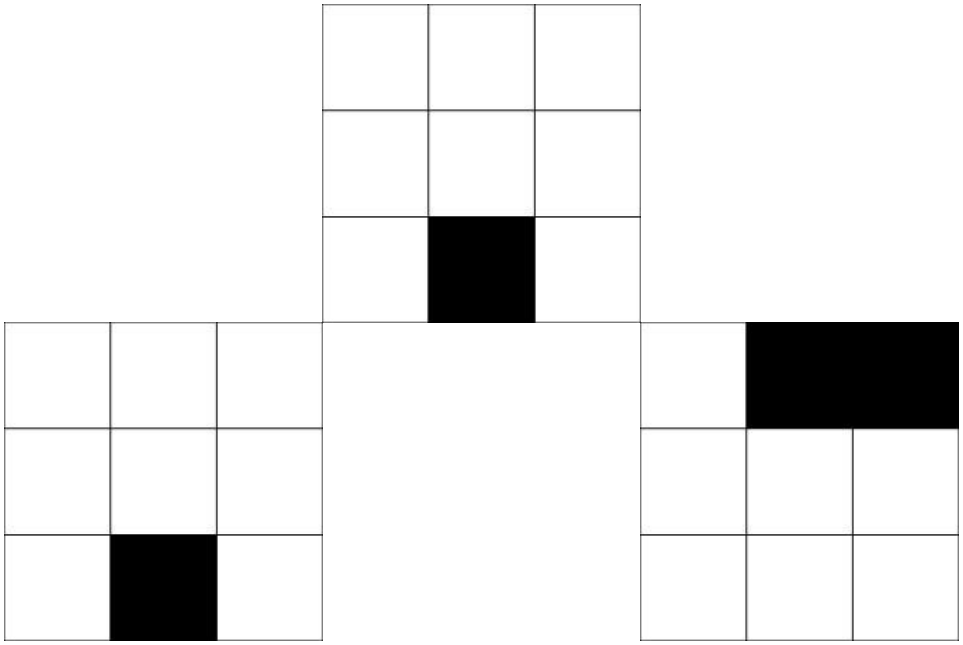
14	62	M	17
15	61	M	17
16	61	F	18
17	59	M	18
18	50	M	13
19	62	M	17
20	62	M	18
21	55	M	14
22	59	M	17
23	65	M	19
24	66	M	11
25	63	M	17
26	69	M	20
27	58	F	17
28	67	M	16
29	62	M	11
30	73	M	22

Supplementary Table S4.1: Demographic and clinical variables in participants with post-stroke aphasia and controls

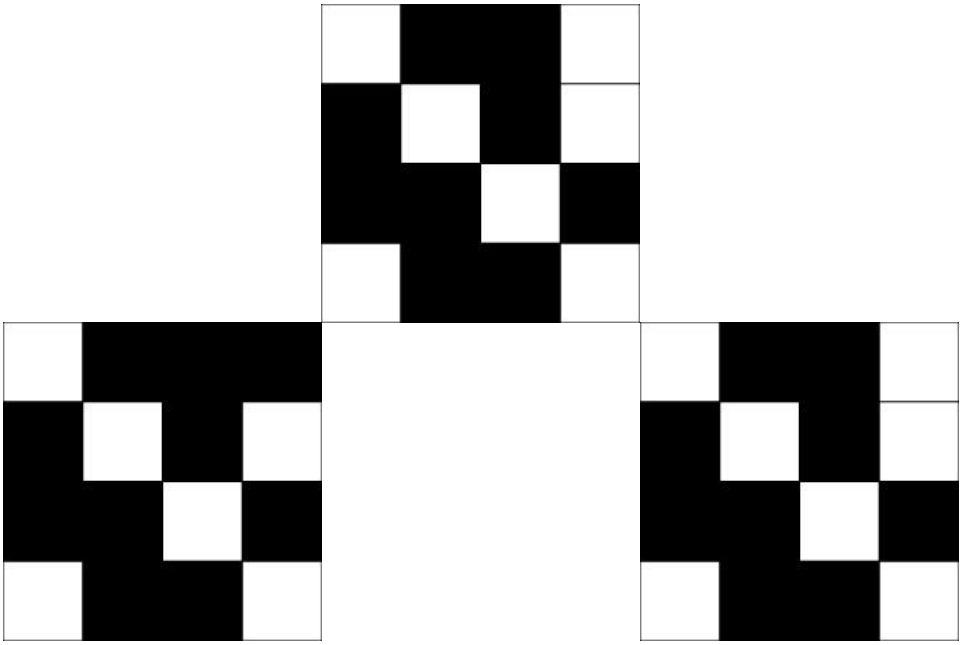
High Cloze Sentences	Low Cloze Sentences
The pilot flies the plane	The bike hits the plane
The gardener mows the lawn	The rocket reaches the lawn
The mother flushes the toilet	The artist paints the toilet
The sunglasses protect her eyes	The chicken pecks her eyes
The secretary answers the phone	The boxer punches the phone
The baker toasts the bread	The blood soaks the bread
The clown juggles the balls	The author signs the balls
The writer sharpens the pencil	The chemist packs the pencil
The boy stamps his feet	The oven burns the feet
The soldier detonates the bomb	The baby shakes the bomb
The referee blows the whistle	The rockstar plays the whistle
The customer leaves a tip	The fiancee buys a tip
The footballer scores the goal	The lion hunts its goal
The burglar picks the lock	The fridge cools the lock
The cat chases the mouse	The nurse dresses the mouse
The dentist pulls out the tooth	The miner digs out the tooth
The cyclist wears the helmet	The snake eats the helmet
The doctor saves the patient	The dog chases the patient
The cleaner sweeps the floor	The teacher lectures the floor
The singer loses her voice	The beer damages her voice
The swimmer gasps for air	The fish loves the air
The jockey rides the horse	The captain orders the horse
The driver brakes the car	The plumber unblocks the car
The pig rolls in the mud	The drummer plays in the mud
The cook washes her hands	The diner chews her hands
The actor forgets his line	The shop closes his line
The bird builds the nest	The bus transports the nest
The waiter clears the table	The butcher carves the table
The sailor scrubs the deck	The boy inherits the deck
The builder lays the brick	The student raises his brick
The fisherman drowns at sea	The camel sleeps at sea
The bartender pours the drink	The mechanic fixes the drink
The fireman climbs the ladder	The scientist discovers the ladder
The stylist combs her hair	The business sells her hair
The politician gives a speech	The toddler babbles a speech
The witch casts the spell	The soldier fires the spell
The maid milks the cow	The lifeguard rescues the cow
The judge clears his throat	The chef slices his throat
The banker counts the money	The farmer sows the money
The runner wins the race	The postman delivers the race

Supplementary Table S4.2: Sentences presented during the 'listening' MRI task

Easy:



Hard:



Supplementary Fig. S4.1: Example easy and hard trials during the 'pattern matching' MRI task

ID	Immediate Word Repetition	Immediate Non-word Repetition	CSB Picture Naming Test	Boston Naming Test	CSB Word-Picture Matching	Forward Digit Span	Backward Digit Span	Cambridge Synonym Judgement Test	CAT Spoken Sentence Comprehension Test	Raven's Coloured Progressive Matrices	Brixton Spatial Anticipation Test	'Cookie Theft' Number of tokens	'Cookie Theft' Words Per Minute	'Cookie Theft' Mean Length of Utterance
Maximum score:	40	30	32	30	32	8	7	48	32	36	55			
Post-stroke aphasia														
1	38	24	25	20	32	6	4	44	32	36	50	134	86.5	13.1
2	39	24	31	28	32	7	6	45	29	31	50	72	116.8	7.9
3	40	28	25	27	32	7	4	48	27	33	38	192	86.6	13.2
4	40	28	22	17	29	4	3	40	25	35	51	188	118.7	13.9
5	40	29	32	28	32	7	3	47	28	36	43	248	126.1	14.5
6	40	23	26	16	32	6	4	48	25	34	48	123	82	12.5
7	27	9	18	9	30	2	1	44	18	20	19	89	36.8	7.7
8	40	28	28	23	31	6	6	48	32	34	41	68	90.7	9.0
9	32	24	17	13	31	5	3	31	25	27	34	108	83.1	10.6
10	40	17	29	26	32	6	4	47	24	27	32	47	65.6	8.6
11	26	13	21	18	32	6	4	48	25	33	36	65	43.3	15.2
12	40	27	28	19	31	5	3	37	24	24	38	53	29.4	5.8
13	40	29	32	28	32	7	3	47	28	36	43	248	126.1	14.5
14	35	19	21	19	32	5	2	44	26	30	38	73	57.6	9.8
15	38	25	19	16	31	3	0	43	28	33	31	79	87.8	11.3

16	38	23	32	25	31	5	2	43	30	32	37	95	66.3	8.7
19	35	17	21	19	31	4	2	48	25	28	41	280	62.9	12.7
20	38	24	31	29	32	4	4	43	27	27	18	39	35.5	6.6
21	33	12	18	16	30	5	2	33	16	34	33	98	38.2	11.3
22	40	27	27	22	32	5	3	42	25	35	52	42	74.1	12.5
24	39	24	31	27	32	7	3	45	22	30	25	220	185.9	13.0
25	39	26	28	17	32	4	2	43	29	31	44	45	60.0	7.4
26	1	0	0	0	19	1	1	44	5	30	41	99	66.7	8.2
27	36	15	25	30	31	4	1	45	20	31	36	117	38.4	11.7
Controls														
1	40	27	31	29	32	8	7	46	30	34	49	217	156.9	12.1
2	40	28	31	28	31	6	3	47	31	34	41	104	178.3	11.4
3	40	29	28	28	32	8	3	47	32	36	38	272	194.3	14.8
4	40	29	31	28	31	8	6	48	32	33	49	105	128.6	7.9
5	40	29	31	30	32	8	6	48	31	34	43	243	169.5	12.1
6	40	29	32	28	32	8	7	48	30	33	51	179	151.3	13.4
7	40	30	32	27	32	8	7	48	31	34	39	685	96.7	16.8
8	40	30	31	30	32	8	7	48	32	36	50	418	128.6	16.9
9	40	29	32	29	32	8	3	48	30	32	42	75	155.2	12.7
10	40	29	30	29	32	7	5	48	32	35	36	105	165.8	13.3
11	40	30	32	27	32	7	6	48	30	36	49	170	122.9	16.7
12	40	28	32	30	32	8	4	48	30	31	42	128	160.0	11.0
13	40	27	32	30	32	8	5	48	32	36	42	185	116.8	9.7
14	40	27	31	26	32	6	4	48	32	32	47	375	202.7	15.3
15	40	28	32	28	32	8	7	48	32	32	45	374	166.2	12.8

16	40	30	31	30	32	8	7	48	30	34	44	159	151.4	13.1
17	39	28	32	28	32	8	4	47	32	34	46	180	116.1	11.9
18	40	29	32	28	32	8	7	47	32	36	51	322	175.6	12.5
19	39	30	30	29	32	5	5	48	30	36	43	129	126.9	13.2
20	40	28	32	29	32	8	7	48	31	34	50	239	140.6	12.3
21	40	27	30	30	32	6	4	46	27	24	45	50	96.8	9.0
22	39	28	31	28	32	8	7	48	31	33	47	237	104.6	14.4
23	40	28	30	28	32	7	4	48	28	30	41	151	161.8	16.6
24	40	29	32	29	32	6	4	48	32	36	39	334	182.2	15.9
25	40	29	32	29	32	8	7	47	32	34	47	157	134.6	15.5
26	40	28	31	30	32	8	7	48	30	36	44	74	116.8	14.5
27	39	28	32	29	32	6	6	48	31	36	48	154	156.6	12.2
28	40	27	32	30	32	6	5	48	29	35	41	203	140.0	11.5
29	40	29	31	29	32	7	6	45	31	33	45	81	211.3	11.3
30	40	27	31	30	32	7	3	48	32	34	47	155	172.2	13.1

Supplementary Table S4.3: Neuropsychological scores in participants with post-stroke aphasia and controls. Neuropsychological scores for the partipants with post-stroke aphasia and controls. Immediate word repetition was from the PALPA9; immediate non-word repetition was from the PALPA8. Abbreviations: CAT = Comprehensive Aphasia Test; CSB = Cambridge Semantic Battery; PALPA = Psycholinguistic Assessments of Language Processing in Aphasia.

Neuropsychological test (maximum)	Maximum score	PSA (mean, SD)	Controls (mean, SD)	P value (uncorrected, FDR corrected)
Immediate Word Repetition	40	35.6 (8.4)	39.9 (0.3)	0.02 / 0.02*
Immediate Non-word Repetition	30	21.5 (7.3)	28.5 (1.0)	0.0001 / 0.0002*
CSB Picture Naming Test	32	24.5 (7.2)	31.2 (0.9)	0.0001 / 0.0002*
Boston Naming Test	30	20.5 (7.2)	28.8 (1.1)	$9x10^{-6}$ / $3x10^{-5}$*
CSB Word-Picture Matching	32	30.9 (2.7)	31.9 (0.3)	0.06 / 0.06
Forward Digit Span	8	5.0 (1.6)	7.3 (0.9)	$4x10^{-7}$ / $2x10^{-6}$*
Backward Digit Span	7	2.9 (1.5)	5.4 (1.5)	$1x10^{-7}$ / $7x10^{-7}$*
Cambridge Synonym Judgement Test	48	43.6 (4.5)	47.6 (0.8)	0.0003 / 0.0005*
CAT Spoken Sentence Comprehension Test	32	24.8 (5.7)	30.8 (1.3)	$3x10^{-5}$ / $8x10^{-5}$*
Raven's Coloured Progressive Matrices	36	31.1 (4.0)	33.8 (2.5)	0.008 / 0.009*
Brixton Spatial Anticipation Test	55	38.3 (9.2)	44.7 (4.1)	0.003 / 0.004*
'Cookie Theft' Number of tokens		117.6 (72.5)	208.7 (131.6)	0.002 / 0.003*
'Cookie Theft' Words Per Minute		77.7 (37.0)	149.4 (30.2)	$1x10^{-9}$ / $1x10^{-8}$*
'Cookie Theft' Mean Length of Utterance		10.8 (2.8)	13.1 (2.3)	0.002 / 0.003*

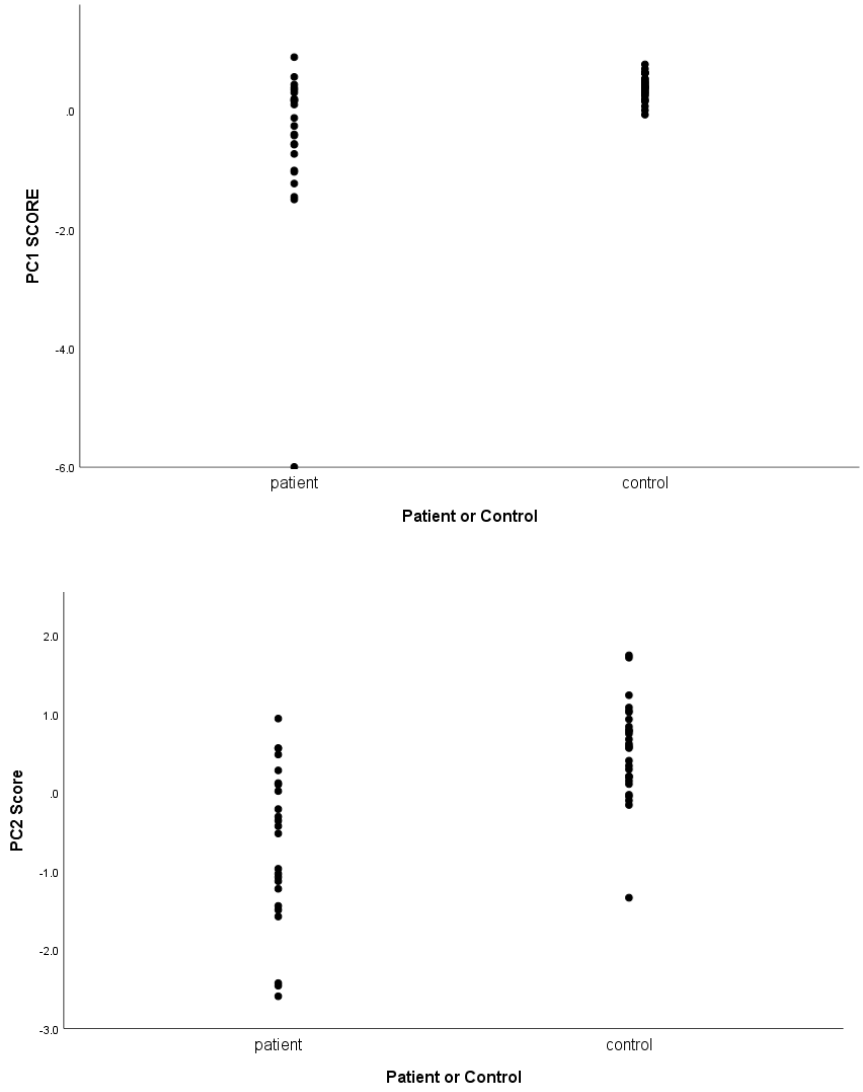
Supplementary Table S4.4: Group level comparisons of neuropsychological scores between participants with post-stroke aphasia and controls. Results of statistical tests comparing neuropsychological scores between the post-stroke aphasia group (n=24) and controls (n=30). Independent samples *t*-tests, not assuming equal variances, were used. Immediate word repetition was from the PALPA9; immediate non-word

repetition was from the PALPA8. * indicates the p-value is significant at q<0.05 (FDR corrected threshold across 14 neuropsychological tests). Abbreviations: CAT = Comprehensive Aphasia Test; CSB = Cambridge Semantic Battery; PALPA = Psycholinguistic Assessments of Language Processing in Aphasia.

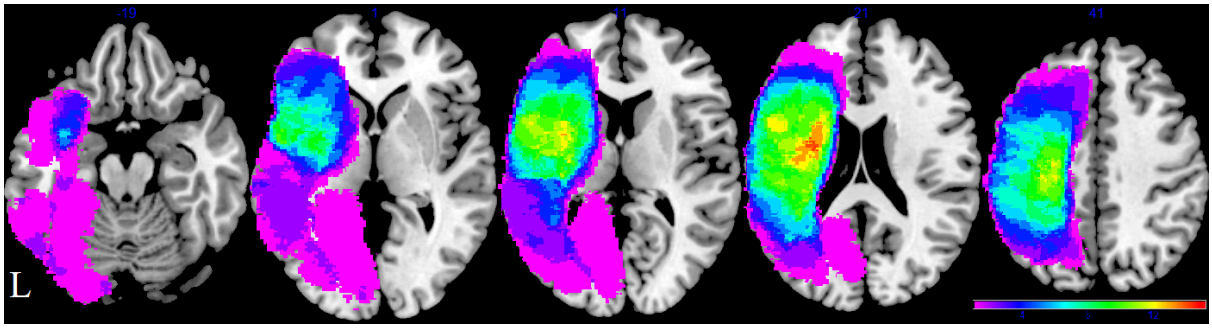
Identifier	PC1	PC2
Post-stroke aphasia		
1	-0.26	0.28
2	0.56	-0.52
3	0.10	0.02
4	-1.03	0.49
5	0.16	0.57
6	-0.40	0.12
7	-1.45	-2.42
8	0.34	-0.35
9	-0.73	-1.49
10	0.38	-1.57
11	-1.23	0.11
12	0.43	-2.46
13	0.16	0.57
14	-0.42	-1.12
15	-0.56	-1.07
16	0.30	-1.22
19	-1.00	-0.21
20	0.89	-2.59
21	-1.50	-1.03
22	-0.13	-0.30
24	0.19	-0.42
25	0.19	-1.44
26	-6.00	0.94
27	-0.57	-0.97
Controls		
1	0.38	0.68
2	0.35	-0.02
3	0.19	0.84
4	0.63	-0.05
5	0.49	0.62
6	0.41	0.78
7	-0.07	1.72
8	0.07	1.75
9	0.64	-0.10
10	0.53	0.21
11	0.16	1.04
12	0.70	-0.15
13	0.63	0.11
14	0.00	1.09
15	0.37	0.94
16	0.53	0.58
17	0.48	0.21

18	0.30	1.24
19	0.26	0.30
20	0.42	0.80
21	0.77	-1.34
22	0.30	0.75
23	0.24	0.34
24	0.15	1.03
25	0.45	0.79
26	0.45	0.57
27	0.33	0.60
28	0.39	0.20
29	0.62	0.15
30	0.39	0.41

Supplementary Table S4.5: Principal Component scores in participants with post-stroke aphasia and controls. Principal Component scores for patients with post-stroke aphasia and controls. Abbreviations: PC = Principal Component.



Supplementary Fig. S4.2: Principal Component scores in participants with post-stroke aphasia and controls



Supplementary Fig. S4.3: Lesion overlap mask of the post-stroke aphasia group. Lesion overlap mask for the 24 patient post-stroke aphasia group in Montreal Neurological Institute space and in neurological convention. Colour bar represents overlap number between 1 and 15. The numbers refer to the Montreal Neurological Institute coordinate space in the Z plane. L = Left.

Cluster	Cluster size (number of voxels)	Peak Coordinate (x y z)	Peak T	Locations
1	3,153	-56, -62, 4	10.52	L Middle Temporal Gyrus L Angular Gyrus, L Superior Temporal Gyrus, L Rolandic Operculum, L Heschl Gyrus, L Insula, L Inferior Parietal, L Postcentral Gyrus, L Supramarginal Gyrus
2	165	-16, 22, 0	5.97	L Caudate Nucleus L Putamen
3	109	-36, -4, -4	5.77	L Insula
4	96	-36, -52, 42	5.64	L Inferior Parietal L Superior Parietal, L Angular Gyrus
5	81	-24, -28, -16	3.94	L Parahippocampal Gyrus
6	58	-54, -60, -14	4.38	L Inferior Temporal Gyrus
7	57	-26, -64, -18	6.33	L Cerebellum L Fusiform Gyrus
8	46	-42, -22, 48	5.40	L Postcentral Gyrus
9	44	-22, -20, 20	5.74	L Thalamus
10	38	-26, 0, -6	5.00	L Putamen
11	34	-42, 36, 24	4.48	L Inferior Frontal Gyrus Triangularis L Middle Frontal Gyrus
12	26	-34, -32, -28	4.43	L Fusiform Gyrus L Cerebellum
13	24	-28, 0, -18	9.13	L Amygdala

Supplementary Table S4.6: Clusters in which lower grey matter tissue integrity was associated with worse language performance (Fig. 4.1). Table showing details of clusters in which lower grey matter tissue integrity was associated with worse language performance, as measured by Principal Component 1 score, in patients and controls combined (n=54), using Voxel Based Correlational Methodology and controlling for lesion volume. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Automated Anatomical Labelling atlas. Abbreviation: L = left; R = right.

Cluster	Cluster size (number of voxels)	Peak Coordinate (x y z)	Peak T	Locations
Positive				
1	329	-51, 37, 1	5.73	L Inferior Frontal Gyrus Triangularis
				L Inferior Frontal Gyrus Orbitalis, L Inferior Frontal Gyrus Opercularis
2	185	54, 37, -2	5.78	R Inferior Frontal Gyrus Orbitalis
				R Inferior Frontal Gyrus Triangularis, R Inferior Frontal Gyrus Opercularis
3	110	21, -77, -38	6.33	R Cerebellum
4	81	-18, -74, -32	6.40	L Cerebellum
5	34	-48, -14, -11	4.11	L Middle Temporal Gyrus
6	27	-57, -44, 1	4.41	L Middle Temporal Gyrus
7	22	-3, 40, 49	4.37	Medial Superior Frontal Gyrus
Negative				
1	1,098	-3, -53, 55	-5.87	Precuneus
				Median Cingulate and Paracingulate Gyri, Posterior Cingulate Gyrus, Cuneus
2	384	0, 40, 13	-6.09	Anterior Cingulate and Paracingulate Gyri
				Median Cingulate and Paracingulate Gyri, Medial Orbital Superior Frontal Gyrus
3	211	54, -2, 1	-6.57	R Superior Temporal Gyrus
				R Rolandic Operculum, R Heschl's Gyrus, R Insula
4	170	-51, 1, 4	-5.83	L Rolandic Operculum
				L Superior Temporal Gyrus, L Insula, L Heschl's Gyrus
5	137	-36, -38, 61	-5.41	L Postcentral Gyrus
				L Superior Parietal Gyrus, L Precentral Gyrus
6	100	51, -53, 4	-4.78	R Middle Temporal Gyrus
				R Inferior Temporal Gyrus
7	49	-27, 40, 37	-4.39	L Dorsolateral Superior Frontal Gyrus
				L Middle Frontal Gyrus

8	42	24, 28, 31	-4.48	R Middle Frontal Gyrus R Dorsolateral Superior Frontal Gyrus
9	31	-18, -11, 70	-4.37	L Precentral Gyrus L Dorsolateral Superior Frontal Gyrus
10	26	42, 7, 10	-4.50	R Insula R Rolandic Operculum, R Inferior Frontal Gyrus Opercularis
11	24	15, -9, 72	-4.27	R Dorsolateral Superior Frontal Gyrus R Supplementary Motor Area
12	22	-40, -82, -2	-4.52	L Middle Occipital Gyrus L Inferior Occipital Gyrus

Supplementary Table S4.7: Clusters of significant sentence cloze activation and deactivation (Fig. 4.2A). Table showing details of clusters of significant 'low>high cloze sentence' activation, using a one-sample *t*-test on the 'univariate parameter estimate' maps of patients and controls combined (n=54). 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Automated Anatomical Labelling atlas. Abbreviation: L = left; R = right.

Cluster	Cluster size (number of voxels)	Peak Coordinate (x y z)	Peak T	Locations
Positive				
1	17,596	30, -95, -5	15.24	R Inferior Occipital Gyrus
				Bilateral: Cerebellum, Fusiform Gyrus, Lingual Gyrus, Inferior Temporal Gyrus, Parahippocampal Gyrus, Hippocampus, Thalamus, Caudate, Putamen, Pallidum, Insula, Inferior Frontal Gyrus Orbitalis, Inferior Frontal Gyrus Triangularis, Inferior Frontal Gyrus Opercularis, Middle Frontal Gyrus Orbitalis, Middle Frontal Gyrus, Precentral Gyrus, Postcentral Gyrus, Inferior Parietal Gyrus, Superior Parietal Gyrus, Angular Gyrus, Precuneus, Cuneus, Supplementary Motor Area, Dorsolateral Superior Frontal Gyrus, Anterior Cingulate and Paracingulate Gyri, Median Cingulate and Paracingulate Gyri, Medial Superior Frontal Gyrus, Supramarginal Gyrus, Rolandic Operculum
2	26	27, -23, 67	4.14	R Precentral Gyrus
				R Postcentral Gyrus
Negative				
1	960	-6, 49, -5	-6.81	Medial Orbital Superior Frontal Gyrus
				Medial Superior Frontal Gyrus, Dorsolateral Superior Frontal Gyrus, Anterior Cingulate and Paracingulate Gyri
2	794	48, -29, 19	-7.60	R Rolandic Operculum
				R Supramarginal Gyrus, R Superior Temporal Gyrus, R Middle Temporal Gyrus, R Temporal Pole Middle Temporal Gyrus, R Temporal Pole Superior Temporal Gyrus, R Insula
3	776	9, -53, 37	-8.50	Precuneus
				Median Cingulate and Paracingulate Gyri, Cuneus, Posterior Cingulate Gyrus
4	639	-45, -65, 43	-7.47	L Angular Gyrus
				L Middle Temporal Gyrus, L Supramarginal Gyrus, L Middle Occipital Gyrus, L Superior Temporal Gyrus, L Insula, L Temporal Pole Superior Temporal Gyrus, L Temporal Pole Middle Temporal Gyrus, L Heschl's Gyrus
5	205	48, -65, 40	-6.16	R Angular Gyrus
				R Middle Temporal Gyrus, R Inferior Parietal Gyrus

6	105	33, -74, -35	-5.97	R Cerebellum
7	63	-36, 31, -17	-4.19	L Inferior Frontal Gyrus Orbitalis L Inferior Frontal Gyrus Triangularis
8	44	-18, -20, -20	-5.40	L Parahippocampal Gyrus L Hippocampus, L Amygdala, L Inferior Frontal Gyrus Orbitalis, L Temporal Pole Superior Temporal Gyrus
9	39	24, -8, -20	-4.09	R Hippocampus R Amygdala, R Parahippocampal Gyrus
10	35	-30, -77, -32	-5.74	L Cerebellum
11	30	27, -50, 13	-5.58	R Calcarine Fissure
12	30	18, -47, 70	-5.30	R Superior Parietal Gyrus
13	29	-33, 22, 46	-4.68	L Middle Frontal Gyrus

Supplementary Table S4.8: Clusters of significant pattern-matching activation and deactivation (Fig. 4.2B). Table showing details of clusters of significant 'hard>easy pattern matching' activation, using a one-sample *t*-test on the 'univariate parameter estimate' maps of patients and controls combined (n=54). 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Automated Anatomical Labelling atlas. Abbreviation: L = left; R = right.

Cluster	Cluster size (number of voxels)	Peak Coordinate (x y z)	Peak T	Locations	Mean sentence cloze decoding All participants (n=54) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)	Mean sentence cloze activation All participants (n=54) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)	Mean pattern matching activation All participants (n=51) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)
1	64	45, 37, -5	6.19	R Inferior Frontal Gyrus Orbitalis R Inferior Frontal Gyrus triangularis, R Middle Frontal Gyrus	4.42 (4.50) t_{53}=7.2, p=3x10^{-9} pFDR=3x10^{-8}*	0.21 (0.45) t_{53}=3.5, p=0.001 pFDR=0.002*	0.06 (0.15) t_{50}=2.9, p=0.006 pFDR=0.01*
2	63	-9, 52, 16	5.38	Medial Superior Frontal Gyrus Anterior Cingulate and Paracingulate Gyri	4.38 (4.95) t_{53}=6.5, p=3x10^{-8} pFDR=9x10^{-8}*	-0.19 (0.53) t_{53}=-2.6, p=0.01 pFDR=0.01*	-0.04 (0.12) t_{50}=-2.3, p=0.03 pFDR=0.05*
3	41	-57, -41, 31	5.07	L Supramarginal Gyrus	4.42 (4.77) t_{53}=6.8, p=9x10^{-9} pFDR=4x10^{-8}*	-0.05 (0.61) t_{53}=-0.7, p=0.51 pFDR=0.51	-0.04 (0.14) t_{50}=-1.8, p=0.08 pFDR=0.09
4	30	23, 50, 19	4.90	R Dorsolateral Superior Frontal Gyrus R Middle Frontal Gyrus, R Medial Superior Frontal Gyrus	4.13 (5.05) t_{53}=6.0, p=2x10^{-7} pFDR=5x10^{-7}*	-0.11 (0.51) t_{53}=-1.6, p=0.11 pFDR=0.12	0.03 (0.09) t_{50}=2.2, p=0.03 pFDR=0.05*
5	29	-45, 40, 4	4.47	L Inferior Frontal Gyrus Triangularis	4.64 (6.44) t_{53}=5.3, p=2x10^{-6} pFDR=3x10^{-6}*	0.31 (0.51) t_{53}=4.5, p=4x10^{-5} pFDR=4x10^{-4}*	0.04 (0.15) t_{50}=1.8, p=0.08 pFDR=0.09
6	28	3, -71, 52	4.69	Precuneus	3.91 (6.25) t_{53}=4.6, p=3x10^{-5} pFDR=3x10^{-5}*	-0.53 (1.05) t_{53}=-3.7, p=5x10^{-4} pFDR=0.002*	0.18 (0.32) t_{50}=4.1, p=2x10^{-4} pFDR=6x10^{-4}*
7	26	3, 31, 55	4.72	Medial Superior Frontal Gyrus	4.46 (6.03)	0.29 (0.61)	0.11 (0.17)

				Supplementary Motor Area	t_{53}=5.4, p=1x10^{-6} pFDR=2x10^{-6}*	t_{53}=3.4, p=0.001 pFDR=0.002*	t_{50}=4.8, p=2x10^{-5} pFDR=9x10^{-5}*
8	25	-24, 25, 43	5.07	L Middle Frontal Gyrus L Dorsolateral Superior Frontal Gyrus	3.81 (6.34) t_{53}=4.4, p=5x10^{-5} pFDR=5x10^{-5}*	-0.23 (0.63) t_{53}=-2.7, p=0.009 pFDR=0.01*	-0.01 (0.14) t_{50}=-0.47, p=0.64 pFDR=0.64
9	24	48,16, 34	4.48	R Inferior Frontal Gyrus Opercularis R Inferior Frontal Gyrus Triangularis	4.91 (6.45) t_{53}=5.6, p=8x10^{-7} pFDR=1x10^{-6}*	0.41 (0.81) t_{53}=3.7, p=5x10^{-4} pFDR=0.002*	0.29 (0.26) t_{50}=8.2, p=8x10^{-11} pFDR=7x10^{-10}*

Supplementary Table S4.9: Clusters of significant sentence cloze decoding (Fig. 4.2C). Table showing details of clusters of significant 'high vs low cloze sentence' decoding information content, using a one-sample *t*-test on the 'multivariate parameter estimate' maps of patients and controls combined (n=54). 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Automated Anatomical Labelling atlas. * indicates the p-value is significant at q<0.05 (FDR corrected threshold across 9 clusters). Abbreviation: L = left; R = right.

Cluster	Cluster size (number of voxels)	Peak Coordinate (x y z)	Peak T	Locations	Mean sentence cloze decoding			
					Patients (n=24) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)	**Controls (n=30) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)**	**ANCOVA Patients vs Controls Controlling ClusterSwc1T1 (F, p-Uncorr) (p-FDR)**	**PC1Score ~ 1 + Group + Mean-PE + Group*Mean-PE + ClusterSwc1T1 All participants (n=54) (B, p-Uncorr for Group*Mean-PE) (p-FDR for Group*Mean-PE)**
All clusters					0.31 (3.93) t_{23}=0.39, p=0.70	3.20 (3.19) t_{29}=5.5, p=0.000006	F(1,51)=5.67, p=0.02	B=-0.15, p=0.002
1	317	-36, 10, 19	7.80	L Inferior Frontal Gyrus Opercularis	0.08 (4.99)	3.02 (4.56)	F(1,51)=2.36, p=0.13	B=-0.13, p=0.0003
				L Inferior Frontal Gyrus Triangularis, L Inferior Frontal Gyrus Orbitalis, L Insula, L Middle Frontal Gyrus, L Putamen, L Precentral Gyrus	t_{23}=0.08, p=0.94 pFDR=0.94	t_{29}=3.6, p=0.001 pFDR=0.002*	pFDR=0.23	pFDR=0.003*
2	127	-18, 49, 28	7.47	L Middle Frontal Gyrus	0.34 (5.58)	4.46 (5.28)	F(1,51)=7.45, p=0.009	B=-0.07, p=0.03
				L Dorsolateral Superior Frontal Gyrus, L Medial Superior Frontal Gyrus	t_{23}=0.3, p=0.77 pFDR=0.94	t_{29}=4.6, p=0.00007 pFDR=0.0005*	pFDR=0.05*	pFDR=0.04*
3	65	-60, -53, -14	6.62	L Inferior Temporal Gyrus	0.92 (5.85)	3.37 (5.87)	F(1,51)=1.25, p=0.27	B=-0.07, p=0.01
				L Middle Temporal Gyrus	t_{23}=0.8, p=0.45 pFDR=0.94	t_{29}=3.1, p=0.004 pFDR=0.007*	pFDR=0.32	pFDR=0.02*

4	55	48, -71, 13	6.65	R Middle Temporal Gyrus	-0.29 (4.59)	3.15 (5.03)	F(1,51)=6.78, p=0.01	B=-0.08, p=0.03
				R Middle Occipital Gyrus	t_{23}=-0.3, p=0.76	t_{29}=3.4, p=0.002	pFDR=0.05*	pFDR=0.04*
					pFDR=0.94	pFDR=0.004*		
5	51	60, -8, 34	5.83	R Postcentral Gyrus	-0.66 (5.50)	1.31 (5.03)	F(1,51)=1.90, p=0.17	B=-0.08, p=0.02
				R Precentral Gyrus	t_{23}=-0.6, p=0.56	t_{29}=1.4, p=0.16	pFDR=0.25	pFDR=0.03*
					pFDR=0.94	pFDR=0.18		
6	45	39, 34, 7	5.57	R Inferior Frontal Gyrus Triangularis	0.54 (4.42)	4.11 (4.76)	F(1,51)=6.70, p=0.01	B=-0.11, p=0.006
				R Insula, R Inferior Frontal Gyrus Opercularis, R Inferior Frontal Gyrus Orbitalis	t_{23}=0.6, p=0.55	t_{29}=4.7, p=0.00005	pFDR=0.05*	pFDR=0.02*
					pFDR=0.94	pFDR=0.0005*		
7	33	18, 61, 10	8.87	R Dorsolateral Superior Frontal Gyrus	-0.11 (6.78)	3.68 (5.67)	F(1,51)=6.13, p=0.02	B=-0.06, p=0.02
				R Medial Superior Frontal Gyrus	t_{23}=-0.1, p=0.94	t_{29}=3.6, p=0.001	pFDR=0.08	pFDR=0.03*
					pFDR=0.94	pFDR=0.002*		
8	31	12, -53, -2	5.66	R Lingual Gyrus	0.84 (6.22)	2.53 (6.04)	F(1,51)=0.99, p=0.32	B=-0.07, p=0.01
				R Precuneus, R Calcarine Fissure	t_{23}=0.7, p=0.51	t_{29}=2.3, p=0.03	pFDR=0.35	pFDR=0.02*
					pFDR=0.94	pFDR=0.04*		
9	30	-54, -71, 4	6.97	L Middle Temporal Gyrus	-0.43 (3.56)	2.42 (5.02)	F(1,51)=3.75, p=0.06	B=-0.14, p=0.001
				L Middle Occipital Gyrus	t_{23}=-0.6, p=0.56	t_{29}=2.6, p=0.01	pFDR=0.18	pFDR=0.005*
					pFDR=0.94	pFDR=0.01*		

10	30	21, -71, 13	5.98	R Calcarine Fissure	-0.34 (8.05)	2.29 (6.01)	F(1,51)=2.50, p=0.12	B=-0.08, p=0.002
				R Cuneus	t_{23}=-0.2, p=0.84 pFDR=0.94	t_{29}=2.1, p=0.05 pFDR=0.06	pFDR=0.23	pFDR=0.008*
11	29	6, -59, 55	5.49	Precuneus	1.48 (6.45) t_{23}=1.1, p=0.27 pFDR=0.94	3.43 (5.97) t_{29}=3.1, p=0.004 pFDR=0.007*	F(1,51)=1.54, p=0.22 pFDR=0.27	B=-0.05, p=0.06 pFDR=0.07
12	28	-11, -62, 52	6.04	L Precuneus	0.59 (6.11) t_{23}=0.5, p=0.64 pFDR=0.94	3.47 (6.27) t_{29}=3.0, p=0.005 pFDR=0.008*	F(1,51)=2.88, p=0.10 pFDR=0.23	B=-0.06, p=0.04 pFDR=0.05*
13	27	54, -35, 13	6.24	R Superior Temporal Gyrus	1.20 (6.82)	3.92 (5.72)	F(1,51)=2.53, p=0.12	B=-0.05, p=0.07
				R Middle Temporal Gyrus	t_{23}=0.9, p=0.40 pFDR=0.94	T29=3.8, p=0.0008 pFDR=0.002*	pFDR=0.23	pFDR=0.08
14	27	3, 31, 52	5.57	Medial Superior Frontal Gyrus	2.74 (5.64)	3.76 (5.72)	F(1,51)=0.23, p=0.63	B=-0.09, p=0.004
				L Dorsolateral Superior Frontal Gyrus	t_{23}=2.4, p=0.03 pFDR=0.63	t_{29}=3.6, p=0.001 pFDR=0.002*	pFDR=0.66	pFDR=0.01*
15	25	-48, -20, -2	6.54	L Middle Temporal Gyrus	0.10 (5.31) t_{23}=0.1, p=0.92 pFDR=0.94	3.36 (6.73) t_{29}=2.7, p=0.01 pFDR=0.01*	F(1,51)=3.87, p=0.05 pFDR=0.18	B=-0.10, p=0.0009 pFDR=0.005*

16	24	-12, 49, 10	6.91	L Anterior Cingulate and Paracingulate Gyri	0.21 (6.31)	4.18 (4.68)	F(1,51)=7.46, p=0.009	B=-0.06, p=0.08
				L Medial Superior Frontal Gyrus	t_{23}=0.2, p=0.87	T29=4.9, p=0.00003	pFDR=0.05*	pFDR=0.08
					pFDR=0.94	pFDR=0.0005*		
17	22	-32, -48, -18	5.32	L Fusiform Gyrus	-0.71 (6.73)	0.98 (5.05)	F(1,51)=1.76, p=0.19	B=-0.06, p=0.04
				L Cerebellum	t_{23}=-0.5, p=0.61	t_{29}=1.1, p=0.30	pFDR=0.25	pFDR=0.05*
					pFDR=0.94	pFDR=0.31		
18	22	-48, 22, -23	5.90	L Temporal Pole Superior Temporal Gyrus	-0.32 (4.03)	0.83 (4.39)	F(1,51)=0.09, p=0.77	B=-0.11, p=0.01
					t_{23}=-0.4, p=0.70	t_{29}=1.0, p=0.31	pFDR=0.77	pFDR=0.02*
					pFDR=0.94	pFDR=0.31		
19	21	-60, -45, 34	6.49	L Supramarginal Gyrus	1.24 (4.02)	4.13 (5.06)	F(1,51)=1.93, p=0.17	B=-0.11, p=0.003
				L Inferior Parietal Gyrus	t_{23}=1.5, p=0.15	t_{29}=4.5, p=0.0001	pFDR=0.25	pFDR=0.01*
					pFDR=0.94	pFDR=0.0005*		
20	20	34, 12, 25	5.25	R Inferior Frontal Gyrus Opercularis	1.28 (5.69)	3.22 (4.64)	F(1,51)=1.82, p=0.18	B=-0.05, p=0.18
				R Rolandic Operculum	t_{23}=1.1, p=0.28	t_{29}=3.8, p=0.0007	pFDR=0.25	pFDR=0.18
					pFDR=0.94	pFDR=0.002*		
21	20	-15, -8, 58	6.13	L Dorsolateral Superior Frontal Gyrus	1.10 (5.16)	3.66 (5.98)	F(1,51)=2.47, p=0.12	B=-0.14, p=0.00002
					t_{23}=1.0, p=0.31	t_{29}=3.3, p=0.002	pFDR=0.23	pFDR=0.0004*
					pFDR=0.94	pFDR=0.004*		

Supplementary Table S4.10: Clusters in which sentence cloze decoding was positively associated with PC1 score – analyses with extracted 'sentence cloze decoding' (Fig. 4.3). Table showing clusters in which 'high vs low cloze sentence' decoding information content was positively associated with out-of-scanner language performance, using a regression of PC1 score on the 'multivariate parameter estimate' maps of patients and controls combined (n=54). Mean 'high vs low' sentence cloze decoding 'multivariate parameter estimate' was extracted from each cluster for each participant and entered into separate t-tests, ANCOVAs and robust regression analyses. 'ClusterSwc1T1' is the mean grey matter tissue integrity of the corresponding cluster. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Automated Anatomical Labelling atlas. * indicates the p-value is significant at $q<0.05$ (FDR corrected threshold across 21 clusters). Abbreviation: B = unstandardised regression coefficient; ANCOVA=Analysis of Covariance; FDR = False Discovery Rate; Mean-PE = mean high vs low cloze sentence decoding ('multivariate parameter estimate'); L = left; R = right.

Cluster	Cluster size (number of voxels)	Peak Coordinate (x y z)	Peak T	Locations	Mean sentence cloze activation				
					Patients (n=24) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)	**Controls (n=30) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)**	**ANOVA Patients vs Controls Controlling ClusterSwc1T1 (F, p-Uncorr) (p-FDR)**	**PC1Score ~ 1 + Mean-PE + ClusterSwc1T1 All participants (n=54) (B, p-Uncorr for Mean-PE) (p-FDR for Mean-PE)**	**PC1Score ~ 1 + Group + Mean-PE + Group*Mean-PE + ClusterSwc1T1 All participants (n=54) (B, p-Uncorr for Group*Mean-PE) (p-FDR for Group*Mean-PE)**
All clusters					0.04 (0.29) t_{23}=0.6, p=0.53	-0.04 (0.28) t_{29}=-0.7, p=0.50	F(1,51)=0.88, p=0.35	B=0.07, p=0.76	B=-0.18, p=0.76
1	317	-36, 10, 19	7.80	L Inferior Frontal Gyrus Opercularis	0.08 (0.27)	0.07 (0.31)	F(1,51)=0.04, p=0.85	B=0.12, p=0.59	B=-0.09, p=0.86
				L Inferior Frontal Gyrus Triangularis, L Inferior Frontal Gyrus Orbitalis, L Insula, L Middle Frontal Gyrus, L Putamen, L Precentral Gyrus	t_{23}=1.5, p=0.16 pFDR=0.37	t_{29}=1.2, p=0.26 pFDR=0.34	pFDR=0.94	pFDR=0.99	pFDR=0.95
2	127	-18, 49, 28	7.47	L Middle Frontal Gyrus	-0.03 (0.46)	-0.03 (0.58)	F(1,51)=0.0002, p=0.99	B=0.07, p=0.60	B=-0.13, p=0.70
				L Dorsolateral Superior Frontal Gyrus, L Medial Superior Frontal Gyrus	t_{23}=-0.3, p=0.77 pFDR=0.95	t_{29}=-0.2, p=0.81 pFDR=0.85	pFDR=0.99	pFDR=0.99	pFDR=0.89
3	65	-60, -53, -14	6.62	L Inferior Temporal Gyrus	0.40 (0.65)	0.35 (0.72)	F(1,51)=0.28, p=0.60	B=0.09, p=0.40	B=-0.27, p=0.28
				L Middle Temporal Gyrus	t_{23}=3.0, p=0.007 pFDR=0.15	t_{29}=2.7, p=0.01 pFDR=0.02 *	pFDR=0.87	pFDR=0.99	pFDR=0.59

4	55	48, -71, 13	6.65	R Middle Temporal Gyrus	-0.21 (0.45)	-0.45 (0.71)	F(1,51)=2.1, p=0.16	B=-0.06, p=0.64	B=1.28, p=0.0003
				R Middle Occipital Gyrus	t_{23}=-2.3, p=0.03	t_{29}=-3.5, p=0.002	pFDR=0.40	pFDR=0.99	pFDR=0.003*
					pFDR=0.17	pFDR=0.006*			
5	51	60, -8, 34	5.83	R Postcentral Gyrus	-0.01 (0.57)	-0.04 (0.51)	F(1,51)=0.04, p=0.84	B=0.01, p=0.94	B=1.64, p=0.000002
				R Precentral Gyrus	t_{23}=-0.1, p=0.90	t_{29}=-0.4, p=0.68	pFDR=0.94	pFDR=0.99	pFDR=0.00004*
					pFDR=0.98	pFDR=0.75			
6	45	39, 34, 7	5.57	R Inferior Frontal Gyrus Triangularis	0.12 (0.39)	0.17 (0.48)	F(1,51)=0.24, p=0.62	B=0.06, p=0.70	B=-0.04, p=0.91
				R Insula, R Inferior Frontal Gyrus Opercularis, R Inferior Frontal Gyrus Orbitalis	t_{23}=1.5, p=0.15	t_{29}=1.9, p=0.07	pFDR=0.87	pFDR=0.99	pFDR=0.96
					pFDR=0.37	pFDR=0.13			
7	33	18, 61, 10	8.87	R Dorsolateral Superior Frontal Gyrus	0.00 (0.66)	-0.18 (0.85)	F(1,51)=1.21, p=0.28	B=0.00, p=0.96	B=0.009, p=0.97
				R Medial Superior Frontal Gyrus	t_{23}=0.03, p=0.98	t_{29}=-1.2, p=0.25	pFDR=0.59	pFDR=0.99	pFDR=0.97
					pFDR=0.98	pFDR=0.34			
8	31	12, -53, -2	5.66	R Lingual Gyrus	-0.23 (0.50)	-0.35 (0.67)	F(1,51)=0.61, p=0.44	B=-0.03, p=0.78	B=0.79, p=0.02
				R Precuneus, R Calcarine Fissure	t_{23}=-2.2, p=0.04	t_{29}=-2.9, p=0.008	pFDR=0.77	pFDR=0.99	pFDR=0.11
					pFDR=0.17	pFDR=0.02*			
9	30	-54, -71, 4	6.97	L Middle Temporal Gyrus	0.00 (0.77)	-0.51 (0.75)	F(1,51)=5.97, p=0.02	B=-0.15, p=0.14	B=0.33, p=0.16
				L Middle Occipital Gyrus	t_{23}=0.02, p=0.98	t_{29}=-3.7, p=0.0008	pFDR=0.40	pFDR=0.99	pFDR=0.42
					pFDR=0.98	pFDR=0.003*			

10	30	21, -71, 13	5.98	R Calcarine Fissure	0.09 (0.55)	-0.16 (0.71)	F(1,51)=2.68, p=0.11	B=-0.02, p=0.86	B=0.15, p=0.63
				R Cuneus	t_{23}=0.8, p=0.42	t_{29}=-1.2, p=0.24	pFDR=0.40	pFDR=0.99	pFDR=0.89
					pFDR=0.72	pFDR=0.34			
11	29	6, -59, 55	5.49	Precuneus	-0.23 (0.53)	-0.53 (0.65)	F(1,51)=3.44, p=0.07	B=-0.18, p=0.15	B=1.06, p=0.0008
					t_{23}=-2.1, p=0.04	t_{29}=-4.5, p=0.0001	pFDR=0.40	pFDR=0.99	pFDR=0.006*
					pFDR=0.17	pFDR=0.002*			
12	28	-11, -62, 52	6.04	L Precuneus	-0.15 (0.44)	-0.38 (0.55)	F(1,51)=2.49, p=0.12	B=0.02, p=0.89	B=-0.36, p=0.31
					t_{23}=-1.7, p=0.11	t_{29}=-3.7, p=0.0008	pFDR=0.40	pFDR=0.99	pFDR=0.59
					pFDR=0.33	pFDR=0.003*			
13	27	54, -35, 13	6.24	R Superior Temporal Gyrus	-0.10 (0.37)	-0.25 (0.34)	F(1,51)=2.29, p=0.14	B=-0.05, p=0.81	B=-0.80, p=0.10
				R Middle Temporal Gyrus	t_{23}=-1.3, p=0.21	t_{29}=-4.0, p=0.0004	pFDR=0.40	pFDR=0.99	pFDR=0.30
					pFDR=0.44	pFDR=0.003*			
14	27	3, 31, 52	5.57	Medial Superior Frontal Gyrus	0.19 (0.49)	0.05 (0.42)	F(1,51)=1.94, p=0.17	B=0.02, p=0.88	B=-0.14, p=0.72
				L Dorsolateral Superior Frontal Gyrus	t_{23}=1.9, p=0.07	t_{29}=0.7, p=0.50	pFDR=0.40	pFDR=0.99	pFDR=0.89
					pFDR=0.25	pFDR=0.62			
15	25	-48, -20, -2	6.54	L Middle Temporal Gyrus	0.22 (0.44)	0.22 (0.43)	F(1,51)=0.005, p=0.95	B=0.03, p=0.85	B=-0.50, p=0.18
					t_{23}=2.5, p=0.02	t_{29}=2.8, p=0.008	pFDR=0.99	pFDR=0.99	pFDR=0.42
					pFDR=0.17	pFDR=0.02*			

16	24	-12, 49, 10	6.91	L Anterior Cingulate and Paracingulate Gyri	-0.08 (0.38)	-0.22 (0.31)	F(1,51)=2.38, p=0.13	B=-0.02, p=0.92	B=-0.22, p=0.66
				L Medial Superior Frontal Gyrus	t_{23}=-1.0, p=0.34	t_{29}=-3.8, p=0.0008	pFDR=0.40	pFDR=0.99	pFDR=0.89
					pFDR=0.65	pFDR=0.003*			
17	22	-32, -48, -18	5.32	L Fusiform Gyrus	0.08 (0.66)	0.07 (0.61)	F(1,51)=0.04, p=0.85	B=0.00, p=0.99	B=-0.57, p=0.03
				L Cerebellum	t_{23}=0.6, p=0.55	t_{29}=0.6, p=0.56	pFDR=0.94	pFDR=0.99	pFDR=0.13
					pFDR=0.72	pFDR=0.65			
18	22	-48, 22, -23	5.90	L Temporal Pole Superior Temporal Gyrus	0.10 (0.75)	0.17 (0.53)	F(1,51)=0.05, p=0.82	B=-0.12, p=0.36	B=0.08, p=0.79
					t_{23}=0.7, p=0.50	t_{29}=1.7, p=0.09	pFDR=0.94	pFDR=0.99	pFDR=0.92
					pFDR=0.72	pFDR=0.16			
19	21	-60, -45, 34	6.49	L Supramarginal Gyrus	0.09 (0.58)	0.00 (0.72)	F(1,51)=0.31, p=0.58	B=0.06, p=0.61	B=-0.48, p=0.05
				L Inferior Parietal Gyrus	t_{23}=0.75, p=0.46	t_{29}=-0.02, p=0.98	pFDR=0.87	pFDR=0.99	pFDR=0.18
					pFDR=0.72	pFDR=0.98			
20	20	34, 12, 25	5.25	R Inferior Frontal Gyrus Opercularis	0.06 (0.47)	0.28 (0.37)	F(1,51)=3.56, p=0.06	B=0.12, p=0.45	B=0.32, p=0.43
				R Rolandic Operculum	t_{23}=0.6, p=0.54	t_{29}=4.2, p=0.0002	pFDR=0.40	pFDR=0.99	pFDR=0.69
					pFDR=0.72	pFDR=0.002*			
21	20	-15, -8, 58	6.13	L Dorsolateral Superior Frontal Gyrus	0.00 (0.21)	-0.07 (0.31)	F(1,51)=0.80, p=0.37	B=-0.07, p=0.78	B=-0.69, p=0.35
					t_{23}=-0.1, p=0.91	t_{29}=-1.3, p=0.21	pFDR=0.71	pFDR=0.99	pFDR=0.61
					pFDR=0.98	pFDR=0.34			

Supplementary Table S4.11: Clusters in which sentence cloze decoding was positively associated with PC1 score – analyses with extracted 'sentence cloze activation' (Fig. 4.3). Table showing clusters in which 'high vs low cloze sentence' decoding information content was positively associated with out-of-scanner language performance, using a regression of PC1 score on the 'multivariate parameter estimate' maps of patients and controls combined (n=54). Mean 'low>high' sentence cloze activation 'univariate parameter estimate' was extracted from each cluster for each participant and entered into separate t-tests, ANCOVAs and robust regression analyses. 'ClusterSwc1T1' is the mean grey matter tissue integrity of the corresponding cluster. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Automated Anatomical Labelling atlas. * indicates the p-value is significant at $q<0.05$ (FDR corrected threshold across 21 clusters). Abbreviation: B = unstandardised regression coefficient; ANCOVA=Analysis of Covariance; FDR = False Discovery Rate; Mean-PE = mean 'low>high' sentence cloze activation ('univariate parameter estimate'); L = left; R = right.

Cluster	Cluster size (number of voxels)	Peak Coordinate (x y z)	Peak T	Locations	Mean pattern matching activation				
					Patients (n=22) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)	**Controls (n=29) (mean, SD) (t-test vs 0, p-Uncorr) (p-FDR)**	**ANOVA Patients vs Controls Controlling ClusterSwc1T1, Response Rate (F, p-Uncorr) (p-FDR)**	**PC1Score ~ 1 + Mean-PE + ClusterSwc1T1 + ResponseRate All participants (n=54) (B, p-Uncorr for Mean-PE) (p-FDR for Mean-PE)**	**PC1Score ~ 1 + Group + Mean-PE + Group*Mean-PE + ClusterSwc1T1 + ResponseRate All participants (n=54) (B, p-Uncorr for Group*Mean-PE) (p-FDR for Group*Mean-PE)**
All clusters					0.03 (0.07) t_{21}=2.2, p=0.04	0.09 (0.08) t_{28}=5.8, p=0.000003	F(1,47)=4.07, p=0.05	B=-0.15, p=0.87	B=1.45, p=0.54
1	317	-36, 10, 19	7.80	L Inferior Frontal Gyrus Opercularis	0.05 (0.08)	0.14 (0.10)	F(1,47)=7.42, p=0.009	B=0.04, p=0.95	B=1.77, p=0.40
				L Inferior Frontal Gyrus Triangularis, L Inferior Frontal Gyrus Orbitalis, L Insula, L Middle Frontal Gyrus, L Putamen, L Precentral Gyrus	t_{21}=2.8, p=0.01 pFDR=0.03*	t_{28}=7.4, $p=4x10^{-8}$ pFDR=$2x10^{-7}$*	pFDR=0.05*	pFDR=0.97	pFDR=0.53
2	127	-18, 49, 28	7.47	L Middle Frontal Gyrus	-0.03 (0.10)	-0.05 (0.11)	F(1,47)=0.03, p=0.87	B=-0.61, p=0.42	B=0.26, p=0.89
				L Dorsolateral Superior Frontal Gyrus, L Medial Superior Frontal Gyrus	t_{21}=-1.3, p=0.19 pFDR=0.25	t_{28}=-2.5, p=0.02 pFDR=0.03*	pFDR=0.90	pFDR=0.95	pFDR=0.89
3	65	-60, -53, -14	6.62	L Inferior Temporal Gyrus	-0.07 (0.12)	-0.05 (0.15)	F(1,47)=0.03, p=0.87	B=-0.26, p=0.61	B=1.80, p=0.17
				L Middle Temporal Gyrus	t_{21}=-2.6, p=0.02	t_{28}=-2.0, p=0.06	pFDR=0.90	pFDR=0.95	pFDR=0.36

					pFDR=0.05*	pFDR=0.07			
4	55	48, -71, 13	6.65	R Middle Temporal Gyrus	0.11 (0.14)	0.28 (0.21)	F(1,47)=8.4, p=0.006	B=0.19, p=0.63	B=0.84, p=0.46
				R Middle Occipital Gyrus	t_{21}=3.7, p=0.001	t_{28}=7.2, p=9x10^{-8}	pFDR=0.05*	pFDR=0.95	pFDR=0.57
					pFDR=0.007*	pFDR=4x10^{-7}*			
5	51	60, -8, 34	5.83	R Postcentral Gyrus	0.05 (0.16)	0.14 (0.16)	F(1,47)=1.70, p=0.20	B=0.20, p=0.68	B=-0.66, p=0.56
				R Precentral Gyrus	t_{21}=1.5, p=0.14	t_{28}=4.8, p=5x10^{-5}	pFDR=0.32	pFDR=0.95	pFDR=0.65
					pFDR=0.22	pFDR=1x10^{-4}*			
6	45	39, 34, 7	5.57	R Inferior Frontal Gyrus Triangularis	0.08 (0.16)	0.17 (0.14)	F(1,47)=3.12, p=0.08	B=0.10, p=0.83	B=1.46, p=0.20
				R Insula, R Inferior Frontal Gyrus Opercularis, R Inferior Frontal Gyrus Orbitalis	t_{21}=2.3, p=0.03	t_{28}=6.3, p=8x10^{-7}	pFDR=0.15	pFDR=0.95	pFDR=0.38
					pFDR=0.06	pFDR=2x10^{-6}*			
7	33	18, 61, 10	8.87	R Dorsolateral Superior Frontal Gyrus	0.05 (0.13)	-0.07 (0.12)	F(1,47)=4.4, p=0.04	B=-0.63, p=0.29	B=3.95, p=0.007
				R Medial Superior Frontal Gyrus	t_{21}=1.7, p=0.10	t_{28}=-2.9, p=0.008	pFDR=0.09	pFDR=0.95	pFDR=0.14
					pFDR=0.18	pFDR=0.01*			
8	31	12, -53, -2	5.66	R Lingual Gyrus	0.00 (0.11)	0.08 (0.13)	F(1,47)=3.75, p=0.06	B=0.59, p=0.38	B=-3.38, p=0.04
				R Precuneus, R Calcarine Fissure	t_{21}=-0.2, p=0.83	t_{28}=3.3, p=0.003	pFDR=0.13	pFDR=0.95	pFDR=0.17
					pFDR=0.83	pFDR=0.005*			
9	30	-54, -71, 4	6.97	L Middle Temporal Gyrus	0.22 (0.23)	0.33 (0.20)	F(1,47)=6.04, p=0.02	B=-0.27, p=0.43	B=1.14, p=0.16

				L Middle Occipital Gyrus	t_{21}=4.4, p=2x10^{-4} pFDR=0.004*	t_{28}=8.8, p=2x10^{-9} pFDR=2x10^{-8}*	pFDR=0.07	pFDR=0.95	pFDR=0.36
10	30	21, -71, 13	5.98	R Calcarine Fissure	0.05 (0.08)	0.12 (0.10)	F(1,47)=7.69, p=0.008	B=0.03, p=0.97	B=4.26, p=0.04
				R Cuneus	t_{21}=2.8, p=0.01 pFDR=0.03*	t_{28}=6.4, p=7x10^{-7} pFDR=2x10^{-6}*	pFDR=0.05*	pFDR=0.97	pFDR=0.17
11	29	6, -59, 55	5.49	Precuneus	0.07 (0.13) t_{21}=2.7, p=0.01 pFDR=0.03*	0.17 (0.14) t_{28}=6.4, p=6x10^{-7} pFDR=2x10^{-6}*	F(1,47)=5.24, p=0.03 pFDR=0.09	B=0.80, p=0.14 pFDR=0.95	B=-0.45, p=0.74 pFDR=0.78
12	28	-11, -62, 52	6.04	L Precuneus	0.03 (0.09) t_{21}=1.5, p=0.16 pFDR=0.22	0.14 (0.16) t_{28}=4.6, p=8x10^{-5} pFDR=2x10^{-4}*	F(1,47)=8.74, p=0.005 pFDR=0.05*	B=0.43, p=0.41 pFDR=0.95	B=-1.93, p=0.22 pFDR=0.39
13	27	54, -35, 13	6.24	R Superior Temporal Gyrus	-0.04 (0.08)	-0.05 (0.07)	F(1,47)=0.02, p=0.90	B=0.75, p=0.46	B=-3.35, p=0.12
				R Middle Temporal Gyrus	t_{21}=-2.7, p=0.01 pFDR=0.03*	t_{28}=-3.4, p=0.002 pFDR=0.004*	pFDR=0.90	pFDR=0.95	pFDR=0.32
14	27	3, 31, 52	5.57	Medial Superior Frontal Gyrus	-0.01 (0.09)	-0.04 (0.10)	F(1,47)=0.77, p=0.38	B=-0.89, p=0.27	B=1.89, p=0.33
				L Dorsolateral Superior Frontal Gyrus	t_{21}=-0.6, p=0.52 pFDR=0.57	t_{28}=-2.5, p=0.02 pFDR=0.03*	pFDR=0.53	pFDR=0.95	pFDR=0.49

15	25	-48, -20, -2	6.54	L Middle Temporal Gyrus	-0.03 (0.09) t_{21}=-1.5, p=0.16 pFDR=0.22	-0.02 (0.08) t_{28}=-1.7, p=0.10 pFDR=0.11	F(1,47)=0.04, p=0.85 pFDR=0.90	B=0.16, p=0.86 pFDR=0.95	B=-4.62, p=0.02 pFDR=0.14
16	24	-12, 49, 10	6.91	L Anterior Cingulate and Paracingulate Gyri L Medial Superior Frontal Gyrus	0.00 (0.06) t_{21}=-0.5, p=0.64 pFDR=0.67	-0.03 (0.07) t_{28}=-2.2, p=0.03 pFDR=0.04*	F(1,47)=2.38, p=0.13 pFDR=0.23	B=-1.11, p=0.30 pFDR=0.95	B=3.05, p=0.24 pFDR=0.39
17	22	-32, -48, -18	5.32	L Fusiform Gyrus L Cerebellum	0.12 (0.14) t_{21}=4.0, p=7x10^{-4} pFDR=0.007*	0.25 (0.15) t_{28}=9.1, p=8x10^{-10} pFDR=2x10^{-8}*	F(1,47)=4.50, p=0.04 pFDR=0.09	B=0.11, p=0.84 pFDR=0.95	B=0.64, p=0.62 pFDR=0.69
18	22	-48, 22, -23	5.90	L Temporal Pole Superior Temporal Gyrus	-0.02 (0.13) t_{21}=-0.7, p=0.49 pFDR=0.57	-0.03 (0.10) t_{28}=-1.5, p=0.13 pFDR=0.13	F(1,47)=0.08, p=0.77 pFDR=0.90	B=-1.30, p=0.05 pFDR=0.95	B=2.51, p=0.10 pFDR=0.30
19	21	-60, -45, 34	6.49	L Supramarginal Gyrus L Inferior Parietal Gyrus	-0.02 (0.11) t_{21}=-0.7, p=0.48 pFDR=0.57	-0.06 (0.15) t_{28}=-2.3, p=0.03 pFDR=0.04*	F(1,47)=0.15, p=0.70 pFDR=0.90	B=0.32, p=0.56 pFDR=0.95	B=-2.49, p=0.07 pFDR=0.25
20	20	34, 12, 25	5.25	R Inferior Frontal Gyrus Opercularis R Rolandic Operculum	0.12 (0.17) t_{21}=3.4, p=0.003 pFDR=0.02*	0.29 (0.20) t_{28}=7.7, p=2x10^{-8} pFDR=1x10^{-7}*	F(1,47)=5.59, p=0.02 pFDR=0.07	B=0.12, p=0.76 pFDR=0.95	B=2.35, p=0.02 pFDR=0.14

21	20	-15, -8, 58	6.13	L Dorsolateral Superior Frontal Gyrus	0.04 (0.08) t_{21}=2.2, p=0.04 pFDR=0.08	0.03 (0.06) t_{28}=2.5, p=0.02 pFDR=0.03 *	F(1,47)=0.80, p=0.38 pFDR=0.53	B=0.27, p=0.81 pFDR=0.95	B=-2.50, p=0.35 pFDR=0.49

Supplementary Table S4.12: Clusters in which sentence cloze decoding was positively associated with PC1 score – analyses with extracted 'pattern matching activation' (Fig. 4.3). Table showing clusters in which 'high vs low cloze sentence' decoding information content was positively associated with out-of-scanner language performance, using a regression of PC1 score on the 'multivariate parameter estimate' maps of patients and controls combined (n=54). Mean 'hard>easy' pattern matching activation 'univariate parameter estimate' was extracted from each cluster for each participant and entered into separate t-tests, ANCOVAs and robust regression analyses. 'ClusterSwc1T1' is the mean grey matter tissue integrity of the corresponding cluster. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Automated Anatomical Labelling atlas. * indicates the p-value is significant at q<0.05 (FDR corrected threshold across 21 clusters). Abbreviation: B = unstandardised regression coefficient; ANCOVA=Analysis of Covariance; FDR = False Discovery Rate; Mean-PE = mean 'hard>easy' pattern matching activation ('univariate parameter estimate'); L = left; R = right.

Chapter 5 Supplementary Material

Supplementary Methods

Participants

One of the patients was left handed but had clear language deficits following stroke suggesting pre-morbid left hemisphere dominance for language; three had additional small infarcts (0.5, 1.5 and 3.5cm^3) in the right hemisphere but of the two whose right sided infarcts predated their left sided stroke, there was no history of ongoing clinical deficit. We included patients with any aphasia severity post-stroke, including those with mild deficits, to enable us to capture the full multidimensional profile of post-stroke aphasia. Three patients had NIH Stroke Scale scores of 0 at 2 weeks post-stroke, despite having ongoing language difficulties, in keeping with the low sensitivity of the NIH Stroke Scale at detecting mild aphasia (Grönberg et al., 2021).

Neuropsychological tests

A subset of the neuropsychological measures tested in patients were available in the controls at both timepoints ('Cinderella ICWs Per Second', 'Cinderella Syllables Per Second', 'Cinderella Total ICWs', 'Cinderella NAS', 'Decision Task IES', 'Speech Task Appropriate ICWs minus Inappropriate ICWs', 'Speech Task Syllable Rate', 'Ravens', 'CAT Cognitive', and 'CAT Fluency').

Each Cinderella story assessment of connected speech production was transcribed from audio recordings and analysed to provide a Narrative Aphasia Score (NAS). This was an adaptation of the Comprehensive Aphasia Test (CAT) picture description scoring method (Swinburn et al., 2005), which did not penalise pure articulatory or dysarthric errors. The NAS was calculated using the following formula:

NAS = ((Appropriate – Inappropriate Information Carrying Words)/Number of utterances) + Syntactic variety + Grammatical well-formedness

Statistical tests

All variables were assessed for normality using the Kolmogorov-Smirnov test. Group differences between patients at T1 and controls were performed using independent samples *t*-tests and Mann-Whitney U tests for normally and non-normally distributed variables, respectively. Changes over time between T1 and T2 were assessed using paired *t*-tests and

Wilcoxon signed-rank tests for normally and non-normally distributed variables, respectively.

Principal Components

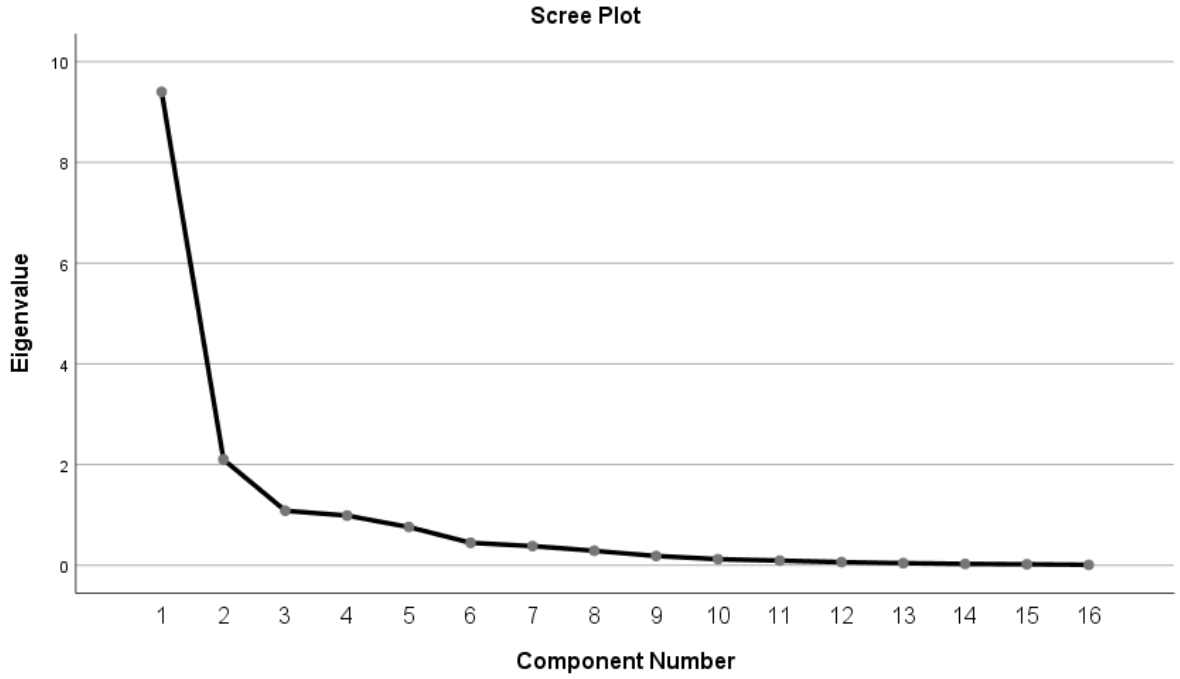
Scores from Principal Components (PCs) with an eigenvalue greater than one were taken to be estimates of underlying cognitive components. To enable direct comparisons to be made between T1 and T2, the patients' scores need to be in the same PCA 'space'. As the T2 data was not appropriate for PCA reduction (see main manuscript), the neuropsychological scores at T2 were back-projected into the 'T1 PCA space' to estimate each patient's PC score had their T2 scores been obtained at T1. This was done by Z-scoring each T2 score using the mean and SD at T1 before multiplying by the component-specific coefficients obtained from the T1 PCA and summating across scores for that participant.

For each PC, a 'lower bound of normal PC score' was calculated, which was taken to be the set of PC scores that would be obtained if a hypothetical individual scored exactly at the fifth percentile of normal on each of the 16 neuropsychological tests. The fifth percentile of normal on each of the 16 neuropsychological tests was estimated using the neuropsychological data obtained from controls in this study, or for tests not performed in the controls of this study, that were published in the CAT manual (Swinburn et al., 2005). The fifth percentile was estimated from control data as:

Fifth percentile = sample mean – (*t*-statistic * sample SD)

Where the *t*-statistic value was for the number of degrees of freedom (n-1) to give a cumulative probability of 0.05 of the control performance being worse than that value.

For each PC, a 'mean PC score' was also calculated, which would have been obtained if an individual scored at the control sample mean on each of the 16 neuropsychological tests at T1. The control sample mean on each of the 16 neuropsychological tests was estimated using the neuropsychological data obtained from controls in this study, or for tests not performed in the controls of this study, that were published in the CAT manual (Swinburn et al., 2005).



Supplementary Figure 5.1: Scree plot from Principal Component Analysis of neuropsychological scores from patients with post-stroke aphasia at Timepoint 1. Scree plot obtained from the Principal Component Analysis of the correlation matrix of neuropsychological test scores from patients with post-stroke aphasia at Timepoint 1 (2 weeks post-stroke). There is a clear inflection point at 3 principal components, indicating that three components should be retained.

Functional MRI data acquisition

A Siemens Magnetom Trio 3T scanner was used to obtain a $1mm^3$ T_1-weighted image with field maps, and a whole-brain T_2*-weighted, gradient echoplanar fMRI sequence consisting of 36 axial slices acquired in an interleaved order (resolution $3.5x3.5x3.0mm^3$; slice thickness 3mm; field of view 225x225x108mm; repetition time 10s; acquisition time 2s; second echo time 31ms; flip angle 90).

Functional MRI preprocessing

EPI volumes were initially preprocessed using FMRIB Software Library (www.fmrib.ox.ac.uk/fsl) (Jenkinson et al., 2012) to perform: motion correction with MCFLIRT (Jenkinson et al., 2002); non-brain removal using BET (Smith, 2002); spatial smoothing using an 8mm full-width at half maximum Gaussian kernel; and highpass temporal filtering with cutoff 100s. Statistical Parametric Mapping (SPM) 12 (Wellcome Centre for Human Neuroimaging, London UK; www.fil.ion.ucl.ac.uk/spm/) was used to: coregister the

mean of run one to the T_1-weighted image in native space before applying this to all EPI volumes; normalise the T_1-weighted image into Montreal Neurological Institute (MNI) space using a modified unified segmentation-normalisation procedure optimised for focally lesioned brain (Seghier et al., 2008); and apply the same normalisation transformation to the coregistered EPI volumes. T_1-weighted images for the same patient were normalised separately at T1 and T2, as structural deformation can occur over time in regions remote from the stroke (Buffon et al., 2005), which could alter the transformation required to map the same patient's image into standard space.

Functional MRI first-level analysis

The canonical HRF without spatial or temporal derivatives was used as it has been validated in sparsely sampled data (Perrachione & Ghosh, 2013) and has been used previously for analysis of sparsely sampled fMRI data during overt speech production (Correia et al., 2020; Simmonds et al., 2014). SPM's FAST pre-whitening option was used to account for temporal autocorrelation (Olszowy et al., 2019).

Functional MRI second-level analysis

A binary lesion mask was manually delineated in each patient's native space at both timepoints by a neurologist (FG). Each lesion mask at T1 was transformed into MNI space using the normalisation transformation obtained previously. ImCalc was used to binarise the overlap of all patients' masks in MNI space and subtract this binarised lesion overlap from the SPM MNI grey matter template. The resultant mask, indicating all grey matter voxels in which no patient had a lesion in MNI space, was used for pre-threshold masking of those patient group second-level analyses which incorporated a language regressor of interest.

Supplementary Results

Identifier	Age	Sex	Years education	Lesion volume (cm^3)	Hours SLT	Time of first scan (days post stroke)	Time of second scan (days post stroke)	NIHSS	
								T1	T2
Post-stroke aphasia									
1	67	M	16	165.6	2	35	90	12	1
2	77	M	14	48.7	2	28	104	4	1
3	68	F	24	49.5	10	30	154	2	0
4	46	F	12	10.3	0	14	119	3	0
5	77	F	21	4.5	0	11	114	2	0
6	50	M	21	54.8	5	12	102	2	0
7	44	M	16	20.5	1.5	12	161	7	1
8	46	M	16	48.7	4	15	200	1	1
9	76	M	25	3.8	0	25	124	2	0
10	60	M	21	12.6	0	10	127	2	1
11	56	M	14	46.5	10	17	96	5	1
12	57	M	18	60.4	7	20	90	6	1
13	75	M	26	36.4	0	16	101	1	0
14	65	M	16	22.7	0	6	101	2	0
15	64	M	22	64.3	7	6	89	13	2
16	64	M	24	25.4	18	12	96	1	0
17	39	F	22	29.9	0	20	91	0	0
18	65	M	14	12.2	0	11	104	5	1
19	49	F	24	5.3	0	18	88	0	0
20	53	M	22	5.6	0	5	102	1	0
21	69	F	15	112.5	4	9	87	4	1
22	54	M	15	35.8	0	14	99	1	0
23	53	F	14	29.7	12	8	92	3	1
24	63	M	18	6.8	0	7	90	1	0
25	50	M	17	9.7	0	20	101	0	0
26	48	F	21	173.9	10	17	95	10	4

Supplementary Table S5.1: Demographic and clinical variables in participants with post-stroke aphasia. Abbreviations: NIHSS = National Institutes of Health Stroke Scale; SLT = Speech-Language Therapy; T1 = timepoint 1; T2 = timepoint 2.

ID	CAT Spoken Picture Description	CAT Fluency	CAT Spoken Comprehension (/66)	CAT Written Comprehension (/62)	CAT Repetition (/74)	CAT Object Naming (/58)	CAT Reading (/70)	CAT Cognitive (/38)	Ravens (/12)	Cinderella ICW Rate (ICWs Per Second)	Cinderella Syllable Rate (Syllables Per Second)	Cinderella Total ICWs	Cinderella Narrative Aphasia Score	Decision Task IES (mean reaction time/proportion correct)	Speech Task Appropriate Minus Inappropriate ICWs	Speech Task Syllable Rate (Syllables Per Second)
Cut-off	20.7	26.1	57.9	55.0	68.8	52.8	58.8	36.7	10.5	0.02	1.76	105.6	21.5	0.50	4.94	1.69
1	42.5	6	47	47	47	48	34	24	11	0.82	2.55	87	9.3	0.55	3.00	1.02
2	60.0	12	56	50	37	44	61	33	11	0.45	1.11	210	9.8	0.47	2.33	1.04
3	38.0	10	58	56	61	55	68	33	7	0.43	1.02	41	9.1	0.57	2.75	1.36
4	69.0	25	64	61	68	57	67	37	11	0.88	2.40	220	13.6	0.41	1.87	0.62
5	108.0	36	63	61	67	57	70	38	12	1.03	2.42	129	16.9	0.50	4.98	1.21
6	34.0	7	55	48	60	43	70	33	12	0.39	1.29	75	12.3	0.47	2.18	1.17
7	0.0	0	52	33	14	6	2	31	9	0.00	0.00	0	0.0	0.32	3.78	1.00
8	44.0	10	58	47	57	45	63	33	10	0.47	1.15	131	13.2	0.39	3.05	0.82
9	61.5	15	60	57	60	57	70	38	9	0.56	1.45	112	17.0	0.40	4.58	1.89
10	40.0	33	62	59	70	57	70	38	12	0.93	2.06	245	23.4	0.41	5.05	2.04
11	31.5	5	35	27	18	14	24	27	4	0.80	1.95	72	7.7	0.63	0.67	1.21
12	22.0	2	60	46	27	26	9	32	11	0.95	2.03	95	7.5	0.37	1.13	0.91
13	115.0	37	66	62	73	57	70	36	12	1.52	3.16	280	29.1	0.38	7.18	2.40
14	31.0	11	58	55	60	47	44	32	8	0.36	0.67	20	4.0	0.35	1.40	1.07
15	26.0	4	54	48	13	20	28	35	11	0.52	1.21	104	9.7	0.42	1.74	0.80
16	106.5	11	46	27	71	47	70	33	12	0.97	1.97	170	18.8	0.37	6.45	2.61

17	73.0	35	64	62	74	56	70	38	12	1.55	3.05	277	24.0	0.39	5.78	1.90
18	52.5	23	62	53	68	54	58	33	12	0.82	2.28	14	9.6	0.40	4.18	1.37
19	87.0	52	66	62	74	58	70	38	12	1.54	3.05	250	24.4	0.36	6.12	2.23
20	122.0	26	55	61	72	58	70	37	12	1.28	2.60	231	21.6	0.45	5.40	1.90
21	63.0	4	32	15	47	23	4	36	6	0.98	1.90	159	13.4	0.89	2.20	0.95
22	25.5	9	52	44	71	44	54	30	7	0.43	0.92	73	9.9	0.49	2.43	1.83
23	11.0	0	48	34	61	35	21	34	8	0.00	0.00	0	0.0	0.51	1.58	0.55
24	62.0	22	65	60	71	57	70	38	11	0.79	1.83	81	17.2	0.37	6.43	2.50
25	68.0	31	64	61	74	58	69	38	12	1.07	2.34	91	14.2	0.33	9.28	3.17
26	-3.0	6	18	2	52	25	8	19	8	0.58	1.19	60	3.3	0.47	-3.65	1.37

Supplementary Table S5.2: Neuropsychological scores in participants with post-stroke aphasia at Timepoint 1. Neuropsychological scores for the 26 participants with post-stroke aphasia at Timepoint 1 (2 weeks post stroke). Values indicate raw test scores except where units are provided in parentheses. The aphasic 'cut-off' was calculated as the fifth percentile of control sample data. Where available, maximum scores are indicated in parentheses. Decision Task Inverse Efficiency Score calculated as mean reaction time (in seconds) divided by the proportion of correct responses; higher Inverse Efficiency Score values indicate worse performance. See Supplementary Methods for description of how Narrative Aphasia Score and 'aphasic cut-off' were calculated. Abbreviations: CAT = Comprehensive Aphasia Test; ICW = Information Carrying Word; ID = Participant Identifier; IES = Inverse Efficiency Score.

ID	CAT Spoken Picture Description	CAT Fluency	CAT Spoken Comprehension (/66)	CAT Written Comprehension (/62)	CAT Repetition (/74)	CAT Object Naming (/58)	CAT Reading (/70)	CAT Cognitive (/38)	Ravens (/12)	Cinderella ICW Rate (ICWs Per Second)	Cinderella Syllable Rate (Syllables Per Second)	Cinderella Total ICWs	Cinderella Narrative Aphasia Score	Decision Task IES (mean reaction time/proportion correct)	Speech Task Appropriate Minus Inappropriate ICWs	Speech Task Syllable Rate (Syllables Per Second)
Cut-off	20.7	26.1	57.9	55.0	68.8	52.8	58.8	36.7	10.5	0.02	1.76	105.6	21.5	0.50	4.94	1.69
1	81.5	12	60	56	52	52	52	36	11	1.57	3.67	172	11.9	0.59	2.08	0.84
2	71.0	23	62	59	64	50	65	38	12	0.86	2.02	216	18.4	0.53	3.20	1.18
3	63.5	24	60	56	74	58	70	38	12	1.00	2.38	167	22.5	0.49	4.93	2.35
4	66.0	57	64	62	72	57	68	37	12	0.92	2.39	179	17.4	0.39	3.58	1.17
5	108.0	36	63	61	67	57	70	38	12	1.27	2.87	324	19.5	0.49	5.22	1.66
6	86.0	30	65	60	67	58	70	37	11	0.90	2.19	249	21.6	0.43	5.47	1.88
7	58.5	28	64	58	73	56	69	37	12	0.00	0.00	0	0.0	0.32	5.65	1.61
8	56.5	16	59	54	73	57	64	32	10	0.71	1.56	143	15.2	0.37	3.43	0.69
9	61.5	23	66	61	70	58	70	38	12	0.97	2.16	227	18.9	0.43	4.40	1.47
10	40.0	25	64	61	70	58	70	38	12	1.10	2.38	256	20.9	0.36	6.08	2.35
11	74.0	12	56	48	46	39	38	34	7	0.96	2.25	174	9.7	0.41	1.82	1.15
12	62.0	27	66	60	66	52	59	37	12	0.89	2.05	145	19.2	0.34	4.33	1.42
13	115.0	37	66	62	73	57	70	36	12	1.52	3.16	280	29.1	0.36	7.08	2.10
14	64.0	16	61	54	68	52	56	37	9	1.09	2.23	189	15.2	0.38	2.27	1.36
15	96.0	31	63	60	65	51	63	38	12	1.03	2.20	191	19.8	0.37	3.97	1.42
16	236.0	34	60	61	74	58	70	38	12	1.41	2.82	271	24.7	0.36	6.48	2.64

17	73.0	50	66	62	74	58	70	38	12	1.59	3.57	300	29.3	0.34	6.38	2.27
18	74.5	28	66	59	73	58	68	37	12	1.25	2.68	129	18.3	0.39	6.02	1.91
19	87.0	52	66	62	74	58	70	38	12	1.54	3.05	250	24.4	0.35	7.59	2.73
20	122.0	43	55	61	72	58	70	37	12	1.26	2.61	243	22.1	0.43	5.03	1.90
21	69.0	20	65	50	63	46	60	38	7	0.90	1.80	240	22.1	0.51	2.99	1.01
22	64.5	17	58	56	69	48	58	32	7	1.50	3.25	90	15.5	0.46	3.53	1.75
23	52.0	15	65	53	63	55	49	34	10	0.70	1.50	108	16.2	0.41	3.45	1.08
24	58.5	41	66	62	74	58	70	38	12	1.41	3.06	160	22.2	0.43	6.08	2.26
25	76.0	35	65	62	74	58	70	38	12	1.39	2.68	151	13.8	0.33	8.93	2.95
26	51.0	10	49	43	63	52	57	36	11	0.56	1.31	87	18.5	0.44	2.48	0.89

Supplementary Table S5.3: Neuropsychological scores in participants with post-stroke aphasia at Timepoint 2. Neuropsychological scores for the 26 participants with post-stroke aphasia at Timepoint 2 (4 months post stroke). Values indicate raw test scores except where units are provided in parentheses. The aphasic 'cut-off' was calculated as the fifth percentile of control sample data. Where available, maximum scores are indicated in parentheses. Decision Task Inverse Efficiency Score calculated as mean reaction time (in seconds) divided by the proportion of correct responses; higher Inverse Efficiency Score values indicate worse performance. See Supplementary Methods for description of how Narrative Aphasia Score and 'aphasic cut-off' were calculated. Abbreviations: CAT = Comprehensive Aphasia Test; ICW = Information Carrying Word; ID = Participant Identifier; IES = Inverse Efficiency Score.

Neuropsychological test	Patients at T1 (median, IQR)	Controls (median, IQR)	P value
CAT Fluency	11.0 (21.5)	46.0 (13.3)	<0.0005*
CAT Cognitive	33.5 (6.0)	38.0 (0.3)	<0.0005*
Ravens	11.0 (4.0)	12.0 (1.0)	0.03
Cinderella Syllables Per Second	1.93 (1.26)	3.29 (1.26)	6×10^{-7}*
Decision Task IES	0.41 (0.12)	0.35 (0.08)	0.004*
Speech Task Appropriate Minus Inappropriate ICWs	3.02 (3.66)	7.16 (1.89)	1×10^{-7}*
Speech Task Syllable Rate	1.29 (0.95)	2.69 (0.65)	<0.0005*

Supplementary Table S5.4: Group level comparisons of neuropsychological scores between participants with post-stroke aphasia at Timepoint 1 and controls. Results of statistical tests comparing neuropsychological scores between the post-stroke aphasia group at Timepoint 1 (n=26) and controls (n=22). 'Cinderella Syllables Per Second' and 'Speech Task Appropriate Minus Inappropriate ICWs' were normally distributed; independent samples *t*-tests were used. 'CAT Fluency', 'CAT Cognitive', 'Ravens', 'Decision Task IES', and 'Speech Task Syllable Rate' were not normally distributed; Mann-Whitney U tests were used. * indicates the p-value is significant at the Bonferroni corrected significance threshold of $p<0.007$ (corrected for 7 comparisons). Abbreviations: CAT = Comprehensive Aphasia Test; ICW = Information Carrying Word; IES = Inverse Efficiency Score; NAS = Narrative Aphasia Score.

Neuropsychological test	Patients at T1 (median, IQR)	Patients at T2 (median, IQR)	P value
CAT Spoken Picture Description	48.3 (40.3)	70.0 (25.5)	<0.0005*
CAT Fluency	11.0 (21.5)	27.5 (19.5)	<0.0005*
CAT Spoken Comprehension	58.0 (12.3)	64.0 (6.0)	<0.0005*
CAT Written Comprehension	51.5 (19.5)	60.0 (5.8)	<0.0005*
CAT Repetition	61.0 (24.0)	70.0 (8.5)	<0.0005*
CAT Object Naming	47.5 (24.3)	57.0 (6.0)	<0.0005*
CAT Reading	65.0 (43.0)	68.5 (11.3)	<0.0005*
CAT Cognitive	33.5 (6.0)	37.0 (2.0)	0.001*
Ravens	11.0 (4.0)	12.0 (1.3)	0.002*
Cinderella ICWs Per Second	0.81 (0.55)	1.06 (0.51)	2×10^{-5}*
Cinderella Syllables Per Second	1.93 (1.26)	2.38 (0.87)	2×10^{-5}*
Cinderella Total ICWs	99.5 (143.5)	184.0 (104.8)	1×10^{-5}*
Cinderella NAS	12.7 (8.8)	19.0 (6.7)	9×10^{-6}*
Decision Task IES	0.41 (0.12)	0.40 (0.09)	0.02
Speech Task Appropriate Minus Inappropriate ICWs	3.02 (3.66)	4.67 (2.71)	0.0005*
Speech Task Syllable Rate	1.29 (0.95)	1.63 (1.10)	0.02

Supplementary Table S5.5: Comparisons of neuropsychological scores between participants with post-stroke aphasia at Timepoint 1 and Timepoint 2. Results of statistical tests comparing neuropsychological scores between the post-stroke aphasia group at Timepoint 1 and Timepoint 2. 'Cinderella ICWs Per Second', 'Cinderella Syllables Per Second', 'Cinderella Total ICWs', 'Cinderella NAS' and 'Speech Task Appropriate Minus Inappropriate ICWs' were normally distributed and used paired *t*-tests. All other variables were not normally distributed and used Wilcoxon signed-rank tests. * indicates the p-value is significant at the Bonferroni corrected significance threshold of $p<0.003$ (corrected for 16 comparisons). Abbreviations: CAT = Comprehensive Aphasia Test; ICW = Information Carrying Word; IES = Inverse Efficiency Score; NAS = Narrative Aphasia Score.

Neuropsychological test	Patients at T2 (median, IQR)	Controls (median, IQR)	P value
CAT Fluency	27.5 (19.5)	46.0 (13.3)	<0.0005*
CAT Cognitive	37.0 (2.0)	38.0 (0.3)	0.01
Ravens	12.0 (1.3)	12.0 (1.0)	0.88
Cinderella Syllables Per Second	2.38 (0.87)	3.29 (1.26)	0.001*
Decision Task IES	0.40 (0.09)	0.35 (0.08)	0.04
Speech Task Appropriate Minus Inappropriate ICWs	4.67 (2.71)	7.16 (1.89)	1×10^{-6}*
Speech Task Syllable Rate	1.63 (1.10)	2.69 (0.65)	<0.0005*

Supplementary Table S5.6: Group level comparisons of neuropsychological scores between participants with post-stroke aphasia at Timepoint 2 and controls. Results of statistical tests comparing neuropsychological scores between the post-stroke aphasia group at Timepoint 2 (n=26) and controls (n=22). 'Cinderella Syllables Per Second' and 'Speech Task Appropriate Minus Inappropriate ICWs' were normally distributed; independent samples *t*-tests were used. 'CAT Fluency', 'CAT Cognitive', 'Ravens', 'Decision Task IES', and 'Speech Task Syllable Rate' were not normally distributed; Mann-Whitney U tests were used. * indicates the p-value is significant at the Bonferroni corrected significance threshold of $p<0.007$ (corrected for 7 comparisons). Abbreviations: CAT = Comprehensive Aphasia Test; ICW = Information Carrying Word; IES = Inverse Efficiency Score; NAS = Narrative Aphasia Score.

Identifier	PC1		PC2		PC3	
	T1	T2	T1	T2	T1	T2
1	0.21	1.54	-0.65	-0.13	-0.44	-0.88
2	-0.04	0.66	0.44	0.44	-0.73	-0.16
3	-1.19	0.51	-0.33	0.11	1.09	0.92
4	0.48	0.61	0.74	0.88	-0.31	0.16
5	0.69	1.68	0.50	0.35	0.32	-0.02
6	-0.74	0.72	0.28	0.52	0.15	0.44
7	-1.40	-2.13	0.75	1.38	-1.97	1.22
8	-0.55	-0.34	0.43	0.37	-0.12	0.35
9	-0.51	0.52	0.42	0.77	0.81	0.15
10	0.57	0.62	0.64	0.78	0.42	0.41
11	0.33	0.48	-1.90	-0.16	-1.03	-0.72
12	0.29	0.24	0.70	1.03	-2.22	-0.07
13	1.83	1.83	0.40	0.52	0.45	0.29
14	-1.53	0.38	0.43	0.24	0.27	0.03
15	0.00	0.78	0.88	0.80	-2.26	-0.12
16	0.46	2.01	-0.65	0.08	1.09	0.89
17	1.55	1.95	0.56	0.64	0.11	0.19
18	-0.42	0.55	0.53	0.63	0.41	0.55
19	1.57	1.52	0.64	0.60	0.40	0.67
20	1.25	1.39	0.08	0.15	0.54	0.53
21	1.20	0.62	-2.69	-0.07	-0.53	-0.03
22	-1.17	0.65	-0.75	-0.64	1.16	0.55
23	-1.85	-0.33	-0.31	0.52	0.03	0.10
24	-0.40	0.99	0.68	0.51	1.15	0.63
25	-0.06	0.52	0.74	0.74	1.38	1.03
26	-0.56	-0.36	-2.56	0.02	-0.17	0.22

Supplementary Table S5.7: Estimated Principal Component scores in participants with post-stroke aphasia. Estimated Principal Component scores for each of the 26 patients with post-stroke aphasia at Timepoint 1 and Timepoint 2. Neuropsychological scores at Timepoint 1 and Timepoint 2 were projected into the Timepoint 1 Principal Component Analysis space. Abbreviations: PC = Principal Component; PC1 = 'fluency' Principal Component; PC2 = 'semantic/executive' principal component; PC3 = 'phonology' principal component; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post stroke).

Cluster	Cluster size (number of voxels)	Coordinate (x y z)	Z	Location
Activation in controls, averaged across timepoints (Fig.5.3A)				
1	44386	-44, -12, 32	7.48	L precentral gyrus
		0, 4, 60	7.36	Supplementary motor cortex
		44, -10, 34	7.35	R precentral gyrus
		62, -18, 0	6.68	R posterior superior temporal gyrus
		54, -6, 22	6.65	R precentral gyrus
		48, -6, 18	6.56	R central opercular cortex
		56, -2, 18	6.54	R precentral gyrus
		16, -16, 0	6.48	R thalamus
		64, -14, 2	6.46	R posterior superior temporal gyrus
		38, -28, 4	6.42	R Heschl's gyrus (includes H1 and H2)
		52, -4, 42	6.35	R precentral gyrus
		60, -4, 38	6.34	R precentral gyrus
		-6, 12, 58	6.26	Superior frontal gyrus
		60, -8, 0	6.21	R planum temporale
		-58, -22, 0	6.13	L posterior superior temporal gyrus
		4, 16, 42	6.12	Paracingulate gyrus
Deactivation in controls, averaged across timepoints (Fig.5.3B)				
1	2405	46, -64, 30	6.38	R superior lateral occipital cortex
		44, -72, 42	5.05	R superior lateral occipital cortex
		62, -56, 18	3.50	R angular gyrus
		64, -54, 14	3.09	R angular gyrus
2	9544	6, -52, 34	6.33	Precuneus
		8, -66, 30	6.14	Precuneus
		-6, -66, 34	5.94	Precuneus
		4, -42, 34	5.92	Posterior cingulate gyrus
		12, -56, 26	5.86	R precuneus
		-10, -60, 20	5.44	Precuneus
		-44, -70, 30	4.84	L superior lateral occipital cortex
		-42, -74, 30	4.78	L superior lateral occipital cortex
		-48, -60, 32	4.42	L angular gyrus
		-24, -42, -4	4.26	L posterior parahippocampal gyrus
		-24, -28, -20	4.23	L posterior parahippocampal gyrus
		4, -32, 46	3.89	Posterior cingulate gyrus

		-26, -46, 4	3.75	L lingual gyrus
		-32, -46, 2	3.68	L temporal occipital fusiform gyrus
		-22, -20, -26	3.51	L anterior parahippocampal gyrus
		-18, -12, -30	3.34	L anterior parahippocampal gyrus
3	6491	-2, 38, -18	5.22	Frontal medial cortex
		-4, 18, -10	4.96	Subcallosal cortex
		4, 18, -10	4.82	Subcallosal cortex
		34, 60, -2	4.78	R frontal pole
		6, 58, -6	4.65	Frontal pole
		8, 14, -14	4.63	Subcallosal cortex
		6, 10, -14	4.61	Subcallosal cortex
		4, 48, -18	4.60	Frontal medial cortex
		6, 24, -12	4.52	Subcallosal cortex
		-2, 8, -14	4.49	Subcallosal cortex
		6, 62, 2	4.42	Frontal pole
		-8, 10, -12	4.40	Subcallosal cortex
		4, 32, -16	4.37	Frontal medial cortex
		-6, 56, -6	4.35	Frontal pole
		-6, 6, -16	4.33	Subcallosal cortex
		-32, 56, -8	4.24	L frontal pole
4	340	24, 32, 38	4.42	R middle frontal gyrus
		18, 44, 46	3.20	R frontal pole
		28, 44, 46	2.72	R frontal pole
		22, 46, 48	2.66	R frontal pole
5	390	-46, 4, -22	3.84	L temporal pole
		-40, -4, -14	3.47	L insular cortex
		-52, 6, -34	3.26	L temporal pole
		-42, -4, 8	3.13	L insular cortex
		-40, -8, -2	3.04	L insular cortex
		-38, 6, -34	2.75	L temporal pole

Supplementary Table S5.8: Clusters of significant activation and deactivation during speech production (Figure 5.3A, 5.3B). Table showing details of local peak maxima in clusters of activation and deactivation in controls, averaged across timepoints 1 and 2. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Harvard-Oxford atlas for cortical regions or the Automated Anatomical Labelling atlas for subcortical regions. Abbreviation: L = left; R = right.

Cluster	Cluster size (number of voxels)	Coordinate (x y z)	Z	Location
Less activation in patients than controls (Fig.5.3C)				
1	3730	24, -54, 18	4.50	R precuneus
		40, -74, 32	3.88	R superior lateral occipital cortex
		46, -66, 30	3.82	R superior lateral occipital cortex
		34, -50, 16	3.71	R posterior supramarginal gyrus
		34, -40, 16	3.64	R planum temporale
		34, -42, 22	3.55	R posterior supramarginal gyrus
		36, -34, 8	3.51	R planum temporale
		20, -60, 42	3.42	R superior lateral occipital cortex
		14, -66, 46	3.41	R precuneus
		20, -64, 44	3.40	R superior lateral occipital cortex
		14, -38, 8	3.28	R posterior cingulate gyrus
		46, -28, -14	3.26	R posterior inferior temporal gyrus
		26, -44, 10	3.23	R posterior cingulate gyrus
		8, -48, 28	3.22	Posterior cingulate gyrus
		52, -54, 24	3.21	R angular gyrus
		54, -56, 36	3.20	R angular gyrus
2	1895	-18, -56, 32	4.09	L precuneus
		-22, -6, 28	3.58	L superior corona radiata
		-20, -38, 34	3.52	L posterior cingulate gyrus
		-20, -20, 20	3.47	L thalamus
		-16, -16, 16	3.45	L thalamus
		-42, -48, 58	3.45	L superior parietal lobule
		-26, -32, 26	3.44	L posterior corona radiata
		-22, -58, 24	3.41	L precuneus
		-20, -8, 22	3.40	L caudate
		-20, -46, 32	3.39	L precuneus
		-16, -4, 16	3.38	L caudate
		-26, -34, 32	3.37	L posterior corona radiata
		-12, -64, 38	3.35	L precuneus
		-22, -24, 30	3.30	L posterior corona radiata
		-10, 10, 2	3.29	L caudate
		-22, -24, 22	3.29	L caudate
3	298	12, -6, -6	3.75	R cerebral peduncle
		4, -6, 8	3.31	Thalamus
		18, -10, 4	3.29	R thalamus
		14, -6, 14	3.08	R thalamus
		18, -2, 0	3.07	R pallidum

Supplementary Table S5.9: Clusters of significantly less activation in patients than controls (Figure 5.3C). Table showing details of local peak maxima in clusters of significantly less activation in patients compared to controls, averaged across Timepoints 1 and 2. This was assessed using an independent sample *t*-

test comparing the ‘mean activation image across timepoints’ of patients vs. controls. ‘Coordinate’ is the Montreal Neurological Institute coordinate of the corresponding peak. ‘Location’ of the peak coordinate is defined using the Harvard-Oxford atlas for cortical regions, the Automated Anatomical Labelling atlas for subcortical regions or the John Hopkins University atlas for white matter tracts. Abbreviation: L = left; R = right.

Cluster	Cluster size (number of voxels)	Coordinate (x y z)	Z	Location
1	3557	42, -36, 4	4.68	R posterior supramarginal gyrus
		60, -56, 4	4.50	R temporooccipital middle temporal gyrus
		56, -54, 6	4.46	R temporooccipital middle temporal gyrus
		42, -44, 12	4.20	R temporooccipital middle temporal gyrus
		58, -18, 20	4.02	R central opercular cortex
		42, -52, 8	3.83	R temporooccipital middle temporal gyrus
		34, -36, 38	3.73	R postcentral gyrus
		32, -42, 36	3.68	R superior parietal lobule
		52, -18, 26	3.50	R postcentral gyrus
		58, -46, 30	3.46	R angular gyrus
		46, -60, 6	3.41	R inferior lateral occipital cortex
		30, -20, 4	3.39	R insular cortex
		62, -42, 8	3.36	R temporooccipital middle temporal gyrus
		34, 2, 14	3.30	R insular cortex
		54, -68, 0	3.29	R inferior lateral occipital cortex
		64, -42, 16	3.26	R posterior supramarginal gyrus

Supplementary Table S5.10: Cluster in which activation was positively associated with fluency at 2 weeks post-stroke (Figure 5.4). Table showing details of local peak maxima for a cluster in which activation was significantly positively associated with Principal Component 1 'fluency' score at Timepoint 1 (2 weeks post-stroke) in patients with post-stroke aphasia. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Harvard-Oxford atlas for cortical regions or the Automated Anatomical Labelling atlas for subcortical regions. Abbreviations: L = left; PC = Principal Component; R = right; T1 = Timepoint 1.

Model/variable	B	SE	p-value	Adjusted R^2	N
Model 1: **T1PC1 ~ 1 + mean activation**			$7.1x10^{-5}$*	0.47	26
Constant	0.15	0.15	0.32		
Mean activation	0.54	0.11	$7.1x10^{-5}$*		
Model 2: **T1PC1 ~ 1 + mean activation + lesion volume + years education + age**			0.003*	0.43	26
Constant	-1.31	1.04	0.22		
Mean activation	0.51	0.12	0.0004*		
Lesion volume	0.002	0.004	0.67		
Years education	0.04	0.04	0.28		
Age	0.009	0.01	0.53		

Supplementary Table S5.11: Regression models for cluster in which activation was positively associated with fluency at 2 weeks post-stroke (Figure 5.4). Activation was positively associated with Principal Component 1 'fluency' score at Timepoint 1 (2 weeks post-stroke) on mass univariate analysis in one cluster. This table contains robust regression models using the mean activation extracted from this cluster to explain speech fluency at 2 weeks post-stroke in patients with post-stroke aphasia. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; PC1 = 'fluency' Principal Component; SE=Standard Error of regression coefficient; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post stroke).

Cluster	Cluster size (number of voxels)	Coordinate (x y z)	Z	Location
1 (Fig. 5.5A)	474	-46, 38, 26	6.66	L frontal pole
		-40, 52, 14	4.19	L frontal pole
		-50, 18, 42	3.58	L middle frontal gyrus
		-42, 18, 50	3.51	L middle frontal gyrus
		-50, 28, 30	3.51	L middle frontal gyrus
		-42, 26, 46	3.44	L middle frontal gyrus
		-46, 18, 46	3.39	L middle frontal gyrus
		-40, 30, 44	3.36	L middle frontal gyrus
		-48, 24, 40	3.31	L middle frontal gyrus
		-32, 60, 14	3.25	L frontal pole
		-38, 58, -2	3.14	L frontal pole
		-46, 30, 34	3.12	L middle frontal gyrus
		-40, 56, -8	3.12	L frontal pole
		-34, 46, 32	3.06	L frontal pole
		-30, 56, 20	2.99	L frontal pole
		-36, 62, -2	2.94	L frontal pole
2 (Fig. 5.5B)	12751	58, -60, 14	4.62	R inferior lateral occipital cortex
		64, -56, -6	4.37	R temporooccipital middle temporal gyrus
		0, -56, 0	4.23	Lingual gyrus
		-42, -64, -14	4.10	L temporal occipital fusiform cortex
		52, -74, 16	4.09	R superior lateral occipital cortex
		-36, -66, -14	4.04	L occipital fusiform gyrus
		30, -46, -28	4.01	R temporal occipital fusiform cortex
		62, -48, 6	4.01	R temporooccipital middle temporal gyrus
		-12, -38, -16	3.98	L posterior parahippocampal gyrus
		-36, -58, -10	3.98	L temporal occipital fusiform cortex
		-44, -60, -16	3.95	L temporal occipital fusiform cortex
		2, -40, -18	3.93	Vermis
		-32, -60, -22	3.91	L temporal occipital fusiform cortex
		-46, -52, -14	3.88	L temporooccipital inferior temporal gyrus
		64, -54, 12	3.86	R temporooccipital middle temporal gyrus
		56, -62, -6	3.85	R inferior lateral occipital cortex
3 (Fig. 5.5C)	469	50, 26, 40	3.90	R middle frontal gyrus
		38, 20, 56	3.81	R middle frontal gyrus
		50, 20, 44	3.72	R middle frontal gyrus
		52, 24, 36	3.70	R middle frontal gyrus
		56, 22, 32	3.56	R middle frontal gyrus
		46, 22, 46	3.43	R middle frontal gyrus
		42, 18, 52	3.41	R middle frontal gyrus
		44, 26, 48	3.40	R middle frontal gyrus
		30, 36, 48	3.34	R frontal pole
		32, 42, 44	3.32	R frontal pole

54, 28, 30	3.23	R middle frontal gyrus
58, 24, 22	3.19	R inferior frontal gyrus pars triangularis
58, 20, 24	3.14	R inferior frontal gyrus pars opercularis
48, 36, 30	3.12	R frontal pole
36, 26, 54	3.11	R middle frontal gyrus
38, 42, 38	3.08	R frontal pole

Supplementary Table S5.12: Regions in which increased activation was positively associated with fluency improvement between 2 weeks and 4 months post-stroke, before controlling for baseline fluency performance (Figure 5.5A-C). Table showing details of local peak maxima for clusters in which increased activation between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) was significantly positively associated with Principal Component 1 'fluency' improvement over the same period in patients with post-stroke aphasia. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Harvard-Oxford atlas for cortical regions or the Automated Anatomical Labelling atlas for subcortical regions. Abbreviations: L = left; PC = Principal Component; R = right; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post-stroke).

Model/variable	B	SE	p-value	Adjusted R^2	N
Cluster 1 (Fig. 5.5A)					
Model 1: **PC1Change ~ 1 + mean activation change**			0.0001*	0.44	26
Constant	0.72	0.12	$2.1x10^{-6}$*		
Mean activation change	0.33	0.07	0.0001*		
Model 2: **PC1Change ~ 1 + mean activation change + T1PC1**			0.0002*	0.48	26
Constant	0.73	0.11	$7.0x10^{-7}$*		
Mean activation change	0.25	0.08	0.003*		
T1PC1	-0.22	0.12	0.09		
(No significant T1PC1*mean activation change interaction)					
Model 3: **PC1Change ~ 1 + mean activation change + T1PC1 + lesion volume + years education + age**			$3.4x10^{-5}$*	0.67	26
Constant	-1.06	0.61	0.10		
Mean activation change	0.29	0.07	0.0003*		
T1PC1	-0.33	0.10	0.006*		
Lesion volume	-0.003	0.002	0.11		
Years education	0.06	0.02	0.02*		
Age	0.01	0.008	0.13		
Cluster 2 (Fig. 5.5B)					
Model 1: **PC1Change ~ 1 + mean activation change**			$5.6x10^{-5}$*	0.48	26
Constant	0.59	0.11	$1.6x10^{-5}$*		
Mean activation change	0.25	0.05	$5.7x10^{-5}$*		
Model 2: **PC1Change ~ 1 + mean activation change + T1PC1**			0.0003*	0.46	26

Constant	0.62	0.11	$9.4x10^{-6}$*		
Mean activation change	0.19	0.06	0.004*		
T1PC1	-0.18	0.13	0.18		
(No significant T1PC1*mean activation change interaction)					
Model 3:			0.005*	0.43	26
PC1Change ~ 1 + mean activation change + T1PC1 + lesion volume + years education + age					
Constant	-0.15	0.89	0.87		
Mean activation change	0.16	0.07	0.04*		
T1PC1	-0.25	0.17	0.15		
Lesion volume	-0.002	0.003	0.45		
Years education	0.008	0.03	0.82		
Age	0.01	0.01	0.28		
Cluster 3 (Fig. 5.5C)					
Model 1:			0.0003*	0.40	26
PC1Change ~ 1 + mean activation change					
Constant	0.69	0.12	$6.0x10^{-6}$*		
Mean activation change	0.30	0.07	0.0003*		
Model 2:			0.0007*	0.42	26
PC1Change ~ 1 + mean activation change + T1PC1					
Constant	0.74	0.11	$1.5x10^{-6}$*		
Mean activation change	0.17	0.08	0.04*		
T1PC1	-0.33	0.13	0.02*		
(No significant T1PC1*mean activation change interaction)					
Model 3:			$3.7x10^{-5}$	0.66	26
PC1Change ~ 1 + mean activation change + T1PC1 + lesion volume + years education + age					
Constant	-1.18	0.62	0.07		
Mean activation change	0.30	0.07	0.0001*		
T1PC1	-0.30	0.11	0.009*		

Lesion volume	-0.005	0.002	0.04*
Years education	0.07	0.02	0.01*
Age	0.01	0.008	0.11

Supplementary Table S5.13: Regression models for clusters in which increased activation was positively associated with fluency improvement between 2 weeks and 4 months post-stroke, before controlling for baseline fluency performance (Figure 5.5A-C). Activation change was positively associated with Principal Component 1 'fluency' score change between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) on mass univariate analysis in three clusters. This table contains robust regression models using the mean activation change extracted from each of these three clusters to explain fluency change between 2 weeks and 4 months post-stroke in patients with post-stroke aphasia. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; PC1 = 'fluency' Principal Component; SE=Standard Error of regression coefficient; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post stroke).

Cluster	Cluster size (number of voxels)	Coordinate (x y z)	Z	Location
1 (Fig. 5.5D)	649	-6, 34, -22	4.47	Frontal medial cortex
		-2, 38, -24	4.16	Frontal medial cortex
		-8, 30, -20	4.06	Subcallosal cortex
		14, 44, -22	3.92	R frontal pole
		0, 38, -30	3.78	Frontal medial cortex
		-12, 10, -26	3.73	L frontal orbital cortex
		6, 36, -26	3.72	Frontal medial cortex
		16, 50, -24	3.62	R frontal pole
		-18, 28, -24	3.57	L frontal orbital cortex
		-4, 12, -26	3.57	Subcallosal cortex
		-8, 14, -26	3.57	Subcallosal cortex
		-16, 32, -24	3.57	L frontal orbital cortex
		-12, 28, -20	3.57	L frontal orbital cortex
		-20, 34, -26	3.55	L frontal orbital cortex
		20, 38, -24	3.54	R frontal pole
		-18, 56, -20	3.47	L frontal pole
2 (Fig. 5.5E)	735	52, -74, 16	4.34	R superior lateral occipital cortex
		58, -60, 14	4.26	R inferior lateral occipital cortex
		56, -68, 14	3.90	R inferior lateral occipital cortex
		68, -36, 32	3.72	R posterior supramarginal gyrus
		64, -54, 14	3.52	R angular gyrus
		64, -56, -6	3.50	R temporooccipital middle temporal gyrus
		60, -60, 22	3.46	R angular gyrus
		62, -62, 2	3.38	R inferior lateral occipital cortex
		64, -50, 18	3.30	R angular gyrus
		60, -62, -2	3.27	R inferior lateral occipital cortex
		68, -44, 12	3.22	R posterior supramarginal gyrus
		66, -46, 26	3.15	R angular gyrus
		60, -50, 12	3.08	R temporooccipital middle temporal gyrus
		68, -42, 20	3.00	R posterior supramarginal gyrus
		62, -48, 6	2.95	R posterior supramarginal gyrus
		58, -66, -2	2.93	R inferior lateral occipital cortex

Supplementary Table S5.14: Regions in which increased activation was positively associated with fluency improvement between 2 weeks and 4 months post-stroke, after controlling for baseline fluency performance (Figure 5.5D-E). Table showing details of local peak maxima for clusters in which increased activation between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) was significantly positively associated with Principal Component 1 'fluency' improvement, controlling for baseline Principal Component 1 score at 2 weeks, in patients with post-stroke aphasia. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Harvard-Oxford atlas for cortical regions or the Automated Anatomical Labelling atlas for

subcortical regions. Abbreviations: L = left; PC = Principal Component; PC1 = ‘fluency’ Principal Component; R = right; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post-stroke).

Model/variable	B	SE	p-value	Adjusted R^2	N
Cluster 1 (Fig. 5.5D)					
Model 1: **PC1Change ~ 1 + mean activation change**			0.02*	0.19	26
Constant	0.72	0.14	$3.3x10^{-5}$*		
Mean activation change	0.32	0.12	0.02*		
Model 2: **PC1Change ~ 1 + mean activation change + T1PC1**			$1.1x10^{-5}$	0.60	26
Constant	0.71	0.10	$1.8x10^{-7}$*		
Mean activation change	0.41	0.09	$8.8x10^{-5}$*		
T1PC1	-0.49	0.10	$5.3x10^{-5}$*		
(No significant T1PC1*mean activation change interaction)					
Model 3: **PC1Change ~ 1 + mean activation change + T1PC1 + lesion volume + years education + age**			$3.5x10^{-5}$*	0.67	26
Constant	-0.48	0.62	0.45		
Mean activation change	0.38	0.08	0.0001*		
T1PC1	-0.57	0.10	$9.5x10^{-6}$*		
Lesion volume	-0.002	0.002	0.33		
Years education	0.04	0.02	0.08		
Age	0.007	0.009	0.40		
Cluster 2 (Fig. 5.5E)					
Model 1: **PC1Change ~ 1 + mean activation change**			$3.3x10^{-5}$*	0.50	26
Constant	0.68	0.11	$1.8x10^{-6}$*		
Mean activation change	0.31	0.06	$3.4x10^{-5}$*		
Model 2: **PC1Change ~ 1 + mean activation change + T1PC1**			$4.4x10^{-5}$*	0.55	26

Constant	0.69	0.10	$7.4x10^{-7}$*		
Mean activation change	0.27	0.06	0.0004*		
T1PC1	-0.20	0.11	0.09		
(No significant T1PC1*mean activation change interaction)					
Model 3:			0.0007*	0.54	26
PC1Change ~ 1 + mean activation change + T1PC1 + lesion volume + years education + age					
Constant	-0.31	0.77	0.69		
Mean activation change	0.23	0.07	0.003*		
T1PC1	-0.28	0.13	0.04*		
Lesion volume	-0.002	0.002	0.44		
Years education	0.03	0.03	0.38		
Age	0.01	0.01	0.33		

Supplementary Table S5.15: Regression models for clusters in which increased activation was positively associated with fluency improvement between 2 weeks and 4 months post-stroke, after controlling for baseline fluency performance (Figure 5.5D-E). Activation change was positively associated with Principal Component 1 'fluency' score change, after controlling for baseline 'fluency' score, between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) on mass univariate analysis in two clusters. This table contains robust regression models using the mean activation change extracted from each of these two clusters to explain fluency change between 2 weeks and 4 months post-stroke in patients with post-stroke aphasia. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; PC1 = 'fluency' Principal Component; SE=Standard Error of regression coefficient; T1 = Timepoint 1.

Cluster	Cluster size (number of voxels)	Coordinate (x y z)	Z	Location
1 (Fig. 5.6A)	1110	-38, 16, -42	4.88	L temporal pole
		18, 40, -20	4.54	R frontal pole
		-40, 18, -38	4.51	L temporal pole
		2, 40, -24	4.36	Frontal medial cortex
		-18, 52, -20	4.35	L frontal pole
		-34, 8, -46	4.31	L temporal pole
		-44, 14, -38	4.24	L temporal pole
		-18, 46, -22	4.22	L frontal pole
		-16, 42, -18	4.17	L frontal pole
		-42, 18, -34	4.14	L temporal pole
		-30, 20, -38	4.08	L temporal pole
		-10, 46, -22	4.04	L frontal pole
		22, 40, -22	4.02	R frontal pole
		-6, 40, -20	4.01	Frontal medial cortex
		-10, 42, -20	4.00	L frontal pole
		-18, 36, -20	3.95	L frontal pole
2 (Fig. 5.6B)	448	44, -14, -40	4.70	R posterior inferior temporal gyrus
		56, -4, -40	4.46	R anterior inferior temporal gyrus
		64, -26, -24	4.27	R posterior inferior temporal gyrus
		56, -32, -24	4.27	R posterior inferior temporal gyrus
		48, -22, -28	4.26	R posterior inferior temporal gyrus
		58, -28, -26	4.26	R posterior inferior temporal gyrus
		58, -24, -28	4.25	R posterior inferior temporal gyrus
		52, -24, -26	4.24	R posterior inferior temporal gyrus
		64, -22, -28	4.18	R posterior inferior temporal gyrus
		56, -10, -40	4.04	R posterior inferior temporal gyrus
		46, -20, -32	3.83	R posterior inferior temporal gyrus
		50, -14, -36	3.59	R posterior inferior temporal gyrus
		56, -2, -36	3.31	R anterior middle temporal gyrus
		56, -40, -22	3.26	R temporooccipital inferior temporal gyrus
		58, -6, -36	3.14	R temporal pole
		58, -12, -34	3.08	R posterior inferior temporal gyrus

Supplementary Table S5.16: Regions in which increased activation was negatively associated with semantic/executive improvement between 2 weeks and 4 months post-stroke, after controlling for baseline semantic/executive score (Figure 5.6). Table showing details of local peak maxima for clusters in which increased activation between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) was significantly negatively associated with Principal Component 2 'semantic/executive' improvement, controlling for Principal Component 2 score at 2 weeks, in patients with post-stroke aphasia. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Harvard-Oxford atlas for cortical regions or the Automated Anatomical Labelling atlas for

subcortical regions. Abbreviations: L = left; PC = Principal Component; R = right; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post-stroke).

Model/variable	B	SE	p-value	Adjusted R^2	N
Cluster 1 (Fig. 5.6A)					
Model 1: **PC2Change ~ 1 + mean activation change**			0.003*	0.28	26
Constant	0.16	0.08	0.06		
Mean activation change	-0.01	0.07	0.84		
Model 2: **PC2Change ~ 1 + mean activation change + T1PC2**			$2.9x10^{-15}$*	0.94	26
Constant	0.39	0.04	$1.8x10^{-10}$*		
Mean activation change	-0.22	0.03	$2.5x10^{-7}$*		
T1PC2	-0.75	0.04	$6.2x10^{-16}$*		
(No significant T1PC2*mean activation change interaction)					
Model 3: **PC2Change ~ 1 + mean activation change + T1PC2 + lesion volume + years education + age**			$4.1x10^{-12}$*	0.94	26
Constant	0.77	0.26	0.008*		
Mean activation change	-0.22	0.03	$3.0x10^{-6}$*		
T1PC2	-0.77	0.06	$1.0x10^{-11}$*		
Lesion volume	-0.0005	0.001	0.66		
Years education	-0.008	0.01	0.43		
Age	-0.003	0.004	0.38		
Cluster 2 (Fig. 5.6B)					
Model 1: **PC2Change ~ 1 + mean activation change**			0.004*	0.27	26
Constant	0.17	0.08	0.05*		
Mean activation change	-0.03	0.10	0.79		
Model 2: **PC2Change ~ 1 + mean activation change + T1PC2**			$6.2x10^{-13}$*	0.91	26

Constant	0.47	0.05	$7.1x10^{-10}$*		
Mean activation change	-0.28	0.06	$8.0x10^{-5}$*		
T1PC2	-0.74	0.05	$1.3x10^{-13}$*		
(No significant T1PC2*mean activation change interaction)					
Model 3:			$6.2x10^{-11}$*	0.91	26
PC2Change ~ 1 + mean activation change + T1PC2 + lesion volume + years education + age					
Constant	1.00	0.30	0.003*		
Mean activation change	-0.29	0.06	$5.2x10^{-5}$*		
T1PC2	-0.71	0.06	$2.9x10^{-10}$*		
Lesion volume	0.0009	0.001	0.49		
Years education	-0.007	0.01	0.52		
Age	-0.007	0.004	0.10		

Supplementary Table S5.17: Regression models for clusters in which increased activation was negatively associated with semantic/executive improvement between 2 weeks and 4 months post-stroke, after controlling for baseline semantic/executive score (Figure 5.6). Activation change was negatively associated with Principal Component 2 'semantic/executive' score change, after controlling for baseline 'semantic/executive' score, between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) on mass univariate analysis in two clusters. This table contains robust regression models using the mean activation change extracted from each of these two clusters to explain semantic/executive change between 2 weeks and 4 months post-stroke in patients with post-stroke aphasia. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; PC2 = 'semantic/executive' principal component; SE=Standard Error of regression coefficient; T1 = Timepoint 1.

Cluster	Cluster size (number of voxels)	Coordinate (x y z)	Z	Location
1 (Fig. 5.7A)	2056	36, 4, 62	5.41	R middle frontal gyrus
		48, 8, 48	4.19	R middle frontal gyrus
		-44, 0, 46	3.91	L precentral gyrus
		4, 38, 50	3.90	Superior frontal gyrus
		52, -28, 52	3.84	R postcentral gyrus
		-38, 2, 56	3.82	L middle frontal gyrus
		-44, 2, 50	3.72	L middle frontal gyrus
		2, 32, 52	3.71	Superior frontal gyrus
		52, -8, 52	3.64	R precentral gyrus
		30, 16, 58	3.63	R middle frontal gyrus
		0, 2, 60	3.59	Supplementary motor cortex
		-30, 8, 60	3.51	L middle frontal gyrus
		12, 34, 52	3.51	R superior frontal gyrus
		-34, 8, 58	3.51	L middle frontal gyrus
		-2, 8, 66	3.50	Supplementary motor cortex
		0, 4, 64	3.48	Supplementary motor cortex
2 (Fig. 5.7B)	1832	-6, 38, -22	4.28	Frontal medial cortex
		-2, 38, -24	4.23	Frontal medial cortex
		-18, 52, -20	4.23	L frontal pole
		12, 44, -24	4.07	R frontal pole
		-18, 44, -24	4.04	L frontal pole
		-14, 64, -10	3.98	L frontal pole
		-16, 48, -22	3.90	L frontal pole
		-18, 28, -22	3.85	L frontal orbital cortex
		14, 48, -24	3.85	R frontal pole
		-16, 24, -22	3.81	L frontal orbital cortex
		14, 64, -4	3.77	R frontal pole
		20, 42, -24	3.75	R frontal pole
		-20, 48, -16	3.74	L frontal pole
		-8, 32, -20	3.71	Frontal medial cortex
		2, 36, -32	3.64	Frontal medial cortex
		4, 40, -32	3.64	Frontal medial cortex
3 (Fig. 5.7C)	819	-40, 56, -2	3.79	L frontal pole
		-48, 38, 22	3.64	L frontal pole
		-24, 44, 28	3.60	L frontal pole
		-42, 52, 6	3.51	L frontal pole
		-40, 44, 20	3.47	L frontal pole
		-32, 58, 0	3.42	L frontal pole
		-40, 56, -8	3.41	L frontal pole
		-40, 54, -14	3.38	L frontal pole
		-46, 38, 26	3.25	L frontal pole
		-20, 54, 18	3.12	L frontal pole

-28, 40, 38	3.12	L frontal pole
-36, 52, -16	3.10	L frontal pole
-34, 44, 8	3.06	L frontal pole
-48, 38, 10	2.80	L frontal pole
-40, 32, 30	2.76	L middle frontal gyrus
-38, 36, 28	2.75	L middle frontal gyrus

Supplementary Table S5.18: Regions in which increased activation was negatively associated with phonology improvement between 2 weeks and 4 months post-stroke, before controlling for baseline phonology score (Figure 5.7A-C). Table showing details of local peak maxima for clusters in which increased activation between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) was significantly negatively associated with Principal Component 3 'phonology' improvement, controlling for Principal Component 3 score at 2 weeks, in patients with post-stroke aphasia. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Harvard-Oxford atlas for cortical regions or the Automated Anatomical Labelling atlas for subcortical regions. Abbreviations: L = left; PC = Principal Component; R = right; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post-stroke).

Model/variable	B	SE	p-value	Adjusted R^2	N
Cluster 1 (Fig. 5.7A)					
Model 1:			$6.4x10^{-5}$*	0.47	26
PC3Change ~ 1 + mean activation change					
Constant	0.08	0.11	0.44		
Mean activation change	-0.27	0.06	$8.3x10^{-5}$*		
Model 2:			$1.3x10^{-12}$*	0.92	26
PC3Change ~ 1 + mean activation change + T1PC3 + T1PC3*mean activation change					
Constant	0.17	0.08	0.04*		
Mean activation change	-0.20	0.07	0.006*		
T1PC3	-0.49	0.10	$6.9x10^{-5}$*		
T1PC3*mean activation change	0.14	0.04	0.0009*		
Model 3:			$3.7x10^{-10}$*	0.91	26
PC3Change ~ 1 + mean activation change + T1PC3 + T1PC3*mean activation change + lesion volume + years education + age					
Constant	0.84	0.50	0.11		
Mean activation change	-0.18	0.07	0.02*		
T1PC3	-0.50	0.13	0.0008*		
T1PC3*mean activation change	0.13	0.04	0.002*		
Lesion volume	-0.0005	0.002	0.79		
Years education	0.002	0.02	0.94		
Age	-0.01	0.007	0.11		
Cluster 2 (Fig. 5.7B)					
Model 1:			0.005*	0.26	26
PC3Change ~ 1 + mean activation change					
Constant	0.10	0.12	0.39		
Mean activation change	-0.19	0.07	0.008*		
Model 2:			$2.7x10^{-9}$*	0.81	26
PC3Change ~ 1 + mean activation change + T1PC3					

Constant	0.23	0.07	0.002*		
Mean activation change	-0.08	0.04	0.09		
T1PC3	-0.60	0.08	$1.2x10^{-7}$*		
Cluster 3 (Fig. 5.7C)					
Model 1:			0.003*	0.28	26
PC3Change ~ 1 + mean activation change					
Constant	0.08	0.12	0.49		
Mean activation change	-0.18	0.06	0.006*		
Model 2:			$3.4x10^{-9}$*	0.80	26
PC3Change ~ 1 + mean activation change + T1PC3					
Constant	0.21	0.07	0.005*		
Mean activation change	-0.10	0.04	0.02*		
T1PC3	-0.57	0.08	$2.1x10^{-7}$*		
(No significant T1PC3*mean activation change interaction)					
Model 3:			$1.7x10^{-7}$*	0.81	26
PC3Change ~ 1 + mean activation change + T1PC3 + lesion volume + years education + age					
Constant	0.88	0.51	0.10		
Mean activation change	-0.11	0.06	0.08		
T1PC3	-0.57	0.12	$8.1x10^{-5}$*		
Lesion volume	-0.0004	0.002	0.84		
Years education	0.0001	0.02	0.99		
Age	-0.01	0.006	0.09		

Supplementary Table S5.19: Regression models for clusters in which increased activation was negatively associated with phonology improvement between 2 weeks and 4 months post-stroke, before controlling for baseline phonology score (Figure 5.7A-C). Activation change was negatively associated with Principal Component 3 'phonology' score change between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) on mass univariate analysis in three clusters. This table contains robust regression models using the mean activation change extracted from each of these three clusters to explain phonology change between 2 weeks and 4 months post-stroke in patients with post-stroke aphasia. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; PC3 = 'phonology' principal component; SE=Standard Error of regression coefficient; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post stroke).

Cluster	Cluster size (number of voxels)	Coordinate (x y z)	Z	Location
1 (Fig. 5.7D)	4088	14, 58, 36	5.59	R frontal pole
		12, 64, 26	5.31	R frontal pole
		14, 62, 32	4.85	R frontal pole
		22, 54, 38	4.62	R frontal pole
		20, 62, 24	4.59	R frontal pole
		26, 48, 42	4.33	R frontal pole
		28, 56, 28	4.21	R frontal pole
		8, 56, 28	4.15	Frontal pole
		-22, 60, 26	4.11	L frontal pole
		14, 54, 42	4.09	R frontal pole
		-16, 46, 36	4.05	L frontal pole
		16, 46, 38	4.04	R frontal pole
		2, 46, 42	4.03	Superior frontal gyrus
		-2, 56, 28	3.99	Superior frontal gyrus
		18, 42, 42	3.94	R frontal pole
		-10, 56, 36	3.86	L frontal pole
2 (Fig. 5.7E)	435	-10, 18, -20	4.27	Subcallosal cortex
		-16, 36, -22	4.20	L frontal pole
		-20, 30, -24	3.97	L frontal orbital cortex
		-16, 24, -22	3.92	L frontal orbital cortex
		0, 38, -26	3.92	Frontal medial cortex
		-8, 28, -18	3.91	Subcallosal cortex
		-6, 30, -22	3.90	Subcallosal cortex
		-12, 34, -20	3.85	L frontal orbital cortex
		20, 38, -24	3.83	R frontal pole
		6, 36, -26	3.82	Frontal medial cortex
		-6, 16, -26	3.78	Subcallosal cortex
		4, 38, -32	3.78	Frontal medial cortex
		-4, 12, -26	3.78	Subcallosal cortex
		-14, 30, -20	3.78	L frontal orbital cortex
		-4, 40, -24	3.73	Frontal medial cortex
		8, 44, -26	3.34	Frontal pole
3 (Fig. 5.7F)	2868	-6, -70, 56	3.98	Precuneus
		-4, -66, 42	3.89	Precuneus
		2, -70, 42	3.74	Precuneus
		8, -64, 64	3.63	Superior lateral occipital cortex
		4, -64, 48	3.62	Precuneus
		0, -88, 32	3.62	Cuneus
		8, -64, 44	3.59	Precuneus
		4, -62, 56	3.51	Precuneus
		4, -86, 36	3.50	Cuneus
		-4, -96, 24	3.49	Occipital pole

-2, -62, 58	3.49	Precuneus
-18, -70, 48	3.36	L superior lateral occipital cortex
-4, -80, 46	3.36	Precuneus
20, -98, 10	3.32	R occipital pole
-8, -80, 48	3.30	Superior lateral occipital cortex
20, -70, 42	3.27	R superior lateral occipital cortex

Supplementary Table S5.20: Regions in which increased activation was negatively associated with phonology improvement between 2 weeks and 4 months post-stroke, after controlling for baseline phonology performance (Figure 5.7D-F). Table showing details of local peak maxima for clusters in which increased activation between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) was significantly negatively associated with Principal Component 3 'phonology' improvement, controlling for Principal Component 3 score at 2 weeks, in patients with post-stroke aphasia. 'Coordinate' is the Montreal Neurological Institute coordinate of the corresponding peak. 'Location' of the peak coordinate is defined using the Harvard-Oxford atlas for cortical regions or the Automated Anatomical Labelling atlas for subcortical regions. Abbreviations: L = left; PC = Principal Component; R = right; T1 = Timepoint 1 (2 weeks post-stroke); T2 = Timepoint 2 (4 months post-stroke).

Model/variable	B	SE	p-value	Adjusted R^2	N
Cluster 1 (Fig. 5.7D)					
Model 1:			0.03*	0.15	26
PC3Change ~ 1 + mean activation change					
Constant	0.04	0.11	0.74		
Mean activation change	-0.05	0.06	0.50		
Model 2:			$1.3x10^{-11}$*	0.90	26
PC3Change ~ 1 + mean activation change + T1PC3 + T1PC3*mean activation change					
Constant	0.24	0.05	0.0001*		
Mean activation change	-0.13	0.03	0.0006*		
T1PC3	-0.65	0.06	$3.7x10^{-10}$*		
T1PC3*mean activation change	0.08	0.04	0.05*		
Model 3:			$1.5x10^{-9}$*	0.90	26
PC3Change ~ 1 + mean activation change + T1PC3 + T1PC3*mean activation change + lesion volume + years education + age					
Constant	0.68	0.40	0.10		
Mean activation change	-0.14	0.05	0.006*		
T1PC3	-0.64	0.08	$1.5x10^{-7}$*		
T1PC3*mean activation change	0.13	0.04	0.01*		
Lesion volume	0.001	0.002	0.54		
Years education	-0.0005	0.02	0.98		
Age	-0.008	0.006	0.19		
Cluster 2 (Fig. 5.7E)					
Model 1:			$1.3x10^{-6}$*	0.61	26
PC3Change ~ 1 + mean activation change					
Constant	0.04	0.10	0.66		
Mean activation change	-0.59	0.09	$1.9x10^{-6}$*		
Model 2:			$5.4x10^{-12}$	0.89	26
PC3Change ~ 1 + mean activation change + T1PC3					

Constant	0.24	0.06	0.0006*		
Mean activation change	-0.32	0.06	$4.2x10^{-5}$*		
T1PC3	-0.68	0.07	$4.6x10^{-10}$*		
(No significant T1PC3*mean activation change interaction)					
Model 3:			$1.1x10^{-10}$*	0.91	26
PC3Change ~ 1 + mean activation change + T1PC3 + lesion volume + years education + age					
Constant	0.47	0.38	0.24		
Mean activation change	-0.28	0.06	0.0001*		
T1PC3	-0.75	0.07	$3.2x10^{-10}$*		
Lesion volume	-0.003	0.001	0.04*		
Years education	0.03	0.01	0.09		
Age	-0.01	0.005	0.08		
Cluster 3 (Fig. 5.7F)					
Model 1:			0.02*	0.17	26
PC3Change ~ 1 + mean activation change					
Constant	0.009	0.11	0.94		
Mean activation change	-0.03	0.06	0.62		
Model 2:			$1.2x10^{-10}$*	0.87	26
PC3Change ~ 1 + mean activation change + T1PC3 + T1PC3*mean activation change					
Constant	0.27	0.06	0.0004*		
Mean activation change	-0.15	0.03	0.0002*		
T1PC3	-0.74	0.07	$8.4x10^{-10}$*		
T1PC3*mean activation change	0.09	0.04	0.03*		
Model 3:			$9.2x10^{-9}$*	0.88	26
PC3Change ~ 1 + mean activation change + T1PC3 + T1PC3*mean activation change + lesion volume + years education + age					
Constant	0.70	0.44	0.13		
Mean activation change	-0.12	0.04	0.007*		
T1PC3	-0.77	0.08	$1.4x10^{-8}$*		
T1PC3*mean activation change	0.07	0.04	0.09		
Lesion volume	-0.001	0.002	0.52		
Years education	0.01	0.02	0.53		

Age	-0.01	0.006	0.13

Supplementary Table S5.21: Regression models for clusters in which increased activation was negatively associated with phonology improvement between 2 weeks and 4 months post-stroke, after controlling for baseline phonology score (Figure 5.7D-F). Activation change was negatively associated with Principal Component 3 'phonology' score change, after controlling for baseline 'phonology' score, between Timepoint 1 (2 weeks) and Timepoint 2 (4 months post-stroke) on mass univariate analysis in three clusters. This table contains robust regression models using the mean activation change extracted from each of these three clusters to explain phonology change between 2 weeks and 4 months post-stroke in patients with post-stroke aphasia. * indicates the p-value is significant at $p<0.05$. Abbreviations: B = unstandardised regression coefficient; N=number of patients included in model; PC = Principal Component; PC3 = 'phonology' principal component; SE=Standard Error of regression coefficient; T1 = Timepoint 1.

References

Alain, C., Du, Y., Bernstein, L. J., Barten, T., & Banai, K. (2018). Listening under difficult conditions: An activation likelihood estimation meta-analysis. *Hum Brain Mapp, 39*(7), 2695-2709.

Allendorfer, J. B., Kissela, B. M., Holland, S. K., & Szaflarski, J. P. (2012). Different patterns of language activation in post-stroke aphasia are detected by overt and covert versions of the verb generation fMRI task. *Medical Science Monitor, 18*(3), CR135-7.

Alyahya, R. S. W., Halai, A., Conroy, P., & Lambon Ralph, M. (2020). A Unified Model of Post-Stroke Language Deficits Including Discourse Production and Their Neural Correlates. *Brain*, 143(5), 1541-1554.

Alyahya, R. S. W., Halai, A. D., Conroy, P., & Lambon Ralph, M. A. (2020). Mapping psycholinguistic features to the neuropsychological and lesion profiles in aphasia. *Cortex*, 124, 260-273.

Anderson, C. A., Wiggins, I. M., Kitterick, P. T., & Hartley, D. E. H. (2017). Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. *Proceedings of the National Academy of Sciences of the United States of America*, 114(38), 10256-10261.

Arthur, W., & Day, D. (1994). Development of a short form for the Raven Advanced Progressive Matrices Test. *Edu Psychol Meas, 54*, 394–403.

Barbieri, E., Mack, J., Chiappetta, B., Europa, E., & Thompson, C. K. (2019). Recovery of offline and online sentence processing in aphasia: Language and domain-general network neuroplasticity. *Cortex*, 120, 394-418.

Barlow, T. (1877). On a case of double cerebral hemiplegia, with cerebral symmetrical lesions. *British Medial Journal*, *2*, 103-104.

Baron, J. C., Levasseur, M., Mazoyer, B., Legault-Demare, F., Mauguière, F., Pappata, S., Tran-Dinh, S. (1992). Thalamocortical diaschisis: positron emission tomography in humans. *J Neurol Neurosurg Psychiatry*, 55(10), 935-942.

Barulli, D., & Stern, Y. (2013). Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. *Trends in Cognitive Sciences, 17*(10), 502-509.

Baumgaertner, A., Hartwigsen, G., & Roman Siebner, H. (2013). Right-hemispheric processing of non-linguistic word features: implications for mapping language recovery after stroke. *Human Brain Mapping*, 34(6), 1293-1305.

Benjamin, E. J., Blaha, M. J., Chiuve, S. E., Cushman, M., Das, S. R., Deo, R., et al. (2017). Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation*, 135(10), e146-e603.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Royal Stat. Soc, 57*, 289–300.

Berthier, M. L., & Pulvermuller, F. (2011). Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nature Reviews Neurology*, 7(2), 86-97.

Berthier, M. L., Pulvermuller, F., Davila, G., Casares, N. G., & Gutierrez, A. (2011). Drug therapy of post-stroke aphasia: A review of current evidence. *Neuropsychology Review, 21*(3), 302-317.

Bestmann, S., Baudewig, J., Siebner, H. R., Rothwell, J. C., & Frahm, J. (2005). BOLD MRI responses to repetitive TMS over human dorsal premotor cortex. *Neuroimage, 28*(1), 22-29.

Binney, R. J., & Lambon Ralph, M. A. (2015). Using a combination of fMRI and anterior temporal lobe rTMS to measure intrinsic and induced activation changes across the

semantic cognition network. *Neuropsychologia*, *76*, 170-181.
Blank, S. C., Bird, H., Turkheimer, F., & Wise, R. J. (2003). Speech production after stroke: the role of the right pars opercularis. *Annals of Neurology*, *54*(3), 310-320.
Boehme, A. K., Martin-Schild, S., Marshall, R. S., & Lazar, R. M. (2016). Effect of aphasia on acute stroke outcomes. *Neurology*, *87*(22), 2348-2354.
Borovsky, A., Saygin, A. P., Bates, E., & Dronkers, N. (2007). Lesion correlates of conversational speech production deficits. *Neuropsychologia*, *45*(11), 2525-2533.
Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, *38*, 1207-1215.
Branzi, F. M., Humphreys, G. F., Hoffman, P., & Lambon Ralph, M. A. (2020). Revealing the neural networks that extract conceptual gestalts from continuously evolving or changing semantic contexts. *Neuroimage*, *220*, 116802.
Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002). *Region of interest analysis using an SPM toolbox* Presented at the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.
Broca, P. (1865). Sur le siège de la facultlé du language articulé. *Bull Soc Anthropol 6*, 337–393.
Brownsett, S. L., Warren, J. E., Geranmayeh, F., Woodhead, Z., Leech, R., & Wise, R. J. (2014). Cognitive control and its impact on recovery from aphasic stroke. *Brain*, 137(1), 242-254.
Bucur, M., & Papagno, C. (2019). Are transcranial brain stimulation effects long-lasting in post-stroke aphasia? A comparative systematic review and meta-analysis on naming performance. *Neuroscience & Biobehavioural Reviews*, 102, 264-289.
Buffon, F., Molko, N., Herve, D., Porcher, R., Denghien, I., Pappata, S., . . . Chabriat, H. (2005). Longitudinal diffusion changes in cerebral hemispheres after MCA infarcts. *Journal of Cerebral Blood Flow and Metabolism*, 25(5), 641-650.
Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton tests*. Thames Valley Test Company.
Burton, H., & McLaren, D. G. (2006). Visual cortex activation in late-onset, Braille naive blind individuals: An fMRI study during semantic and phonological tasks with heard words. *Neuroscience Letters*, 392(1-2), 38-42.
Butler, R. A., Lambon Ralph, M. A., & Woollams, A. M. (2014). Capturing multidimensionality in stroke aphasia: mapping principal behavioural components to neural structures. *Brain*, 137(12), 3248-3266.
Cardebat, D., Demonet, J. F., De Boissezon, X., Marie, N., Marie, R. M., Lambert, J., . . . Study, P. E. T. L. A. (2003). Behavioral and neurofunctional changes over time in healthy and aphasic subjects: a PET Language Activation Study. *Stroke*, *34*(12), 2900-2906.
Carota, F., Kriegeskorte, N., Nili, H., & Pulvermüller, F. (2017). Representational Similarity Mapping of Distributional Semantics in Left Inferior Frontal, Middle Temporal, and Motor Cortex. *Cereb Cortex*, *27*(1), 294-309.
Carrera, E., & Tononi, G. (2014). Diaschisis: past, present, future. *Brain*, *137*(9), 2408-2422.
Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: a meta-analysis. *Learning and Individual Differences 19*(2), 246-251.
Castren, E., & Hen, R. (2013). Neuronal plasticity and antidepressant actions. *Trends in Neurosciences*, *36*(5), 259-267.
Cattell, R. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245-276.
Chalela, J. A., Kidwell, C. S., Nentwich, L. M., Luby, M., Butman, J. A., Demchuk, A. M., . . . Warach, S. (2007). Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: a prospective comparison. *Lancet*,

369(9558), 293-298.
Chang, Y. N., & Lambon Ralph, M. A. (2020). A unified neurocomputational bilateral model of spoken language production in healthy participants and recovery in poststroke aphasia. *Proc Natl Acad Sci U S A, 117*(51), 32779-32790.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale.
Collignon, O., Dormal, G., Albouy, G., Vandewalle, G., Voss, P., Phillips, C., & Lepore, F. (2013). Impact of blindness onset on the functional organization and the connectivity of the occipital cortex. *Brain, 136*, 2769-2783.
Conroy, P., Sotiropoulou Drosopoulou, C., Humphreys, G. F., Halai, A. D., & Lambon Ralph, M. A. (2018). Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia. *Brain, 141*(6), 1815-1827.
Corbetta, M., Kincade, M. J., Lewis, C., Snyder, A. Z., & Sapir, A. (2005). Neural basis and recovery of spatial attention deficits in spatial neglect. *Nature Neuroscience, 8*(11), 1603-1610.
Correia, J. M., Caballero-Gaudes, C., Guediche, S., & Carreiras, M. (2020). Phonatory and articulatory representations of speech production in cortical and subcortical fMRI responses. *Sci Rep, 10*(1), 4529.
Coutanche, M. N., Thompson-Schill, S. L., & Schultz, R. T. (2011). Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity. *Neuroimage, 57*(1), 113-123.
Cox, R. W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res, 29*(3), 162-173.
Cox, R. W., & Hyde, J. S. (1997). Software tools for analysis and visualization of fMRI data. *NMR Biomed, 10*(4-5), 171-178.
Crinion, J., & Price, C. J. (2005). Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. *Brain, 128*(12), 2858-2871.
Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain, 126*(5), 1193-1201.
Crisp, J., & Lambon Ralph, M. A. (2006). Unlocking the nature of the phonological-deep dyslexia continuum: the keys to reading aloud are in phonology and semantics. *J Cogn Neurosci, 18*(3), 348-362.
Crittenden, B. M., Mitchell, D. J., & Duncan, J. (2016). Task Encoding across the Multiple Demand Cortex Is Consistent with a Frontoparietal and Cingulo-Opercular Dual Networks Distinction. *Journal of Neuroscience, 36*(23), 6147-6155.
Davis, C. J. (2005). N-watch: a program for deriving neighborhood size and other psycholinguistic statistics. *Behav Res Methods, 37*(1), 65-70.
Davis, T., LaRocque, K. F., Mumford, J. A., Norman, K. A., Wagner, A. D., & Poldrack, R. A. (2014). What do differences between multi-voxel and univariate analysis mean? How subject-, voxel-, and trial-level variance impact fMRI analysis. *Neuroimage, 97*, 271-283.
Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychol Rev, 104*(4), 801-838.
Demeurisse, G., & Capon, A. (1991). Brain activation during a linguistic task in conduction aphasia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 27*(2), 285-294.
Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage, 31*(3), 968-980.
Di Lazzaro, V., Oliviero, A., Profice, P., Insola, A., Mazzone, P., Tonali, P., & Rothwell, J. (1999). Direct demonstration of interhemispheric inhibition of the human motor cortex

produced by transcranial magnetic stimulation. *Exp Brain Res*, *124*(4), 520-524.
Dignam, J., Copland, D., McKinnon, E., Burfein, P., O'Brien, K., Farrell, A., & Rodriguez, A. D. (2015). Intensive Versus Distributed Aphasia Therapy: A Nonrandomized, Parallel-Group, Dosage-Controlled Study. *Stroke*, *46*(8), 2206-2211.
Dignam, J. K., Rodriguez, A. D., & Copland, D. A. (2016). Evidence for Intensive Aphasia Therapy: Consideration of Theories From Neuroscience and Cognitive Psychology. *PM and R*, *8*(3), 254-267.
Dosenbach, N. U., Fair, D. A., Cohen, A. L., Schlaggar, B. L., & Petersen, S. E. (2008). A dual-networks architecture of top-down control. *Trends Cogn Sci*, *12*(3), 99-105.
Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nat Neurosci*, *3*(3), 277-283.
Duncan, J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. *Trends Cogn Sci*, *14*(4), 172-179.
Dunst, B., Benedek, M., Jauk, E., Bergner, S., Koschutnig, K., Sommer, M., . . . Neubauer, A. C. (2014). Neural efficiency as a function of task demands. *Intelligence*, *42*(100), 22-30.
Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(24), 13763-13768.
Eickhoff, S., & Bzdok, D. (2013). Meta-analyses in basic and clinical neuroscience: state of the art and perspective. In U. S & J. O (Eds.), *fMRI* (pp. 77-87). Springer.
Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *Neuroimage*, *59*(3), 2349-2361.
Eickhoff, S. B., Bzdok, D., Laird, A. R., Roski, C., Caspers, S., Zilles, K., & Fox, P. T. (2011). Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. *Neuroimage*, *57*(3), 938-949.
Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp*, *30*(9), 2907-2926.
Eickhoff, S. B., Nichols, T. E., Laird, A. R., Hoffstaedter, F., Amunts, K., Fox, P. T., . . . Eickhoff, C. R. (2016). Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *Neuroimage*, *137*, 70-85.
El Hachioui, H., Visch-Brink, E. G., de Lau, L. M., van de Sandt-Koenderman, M. W., Nouwens, F., Koudstaal, P. J., & Dippel, D. W. (2017). Screening tests for aphasia in patients with stroke: a systematic review. *J Neurol*, *264*(2), 211-220.
Elkana, O., Frost, R., Kramer, U., Ben-Bashat, D., & Schweiger, A. (2013). Cerebral language reorganization in the chronic stage of recovery: a longitudinal fMRI study. *Cortex*, *49*(1), 71-81.
Ellis, C., Simpson, A. N., Bonilha, H., Mauldin, P. D., & Simpson, K. N. (2012). The one-year attributable cost of poststroke aphasia. *Stroke*, *43*(5), 1429-1431.
Elsner, B., Kugler, J., Pohl, M., & Mehrholz, J. (2015). Transcranial direct current stimulation (tDCS) for improving aphasia in patients with aphasia after stroke. *The Cochrane database of systematic reviews*, 5(CD009760).
Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., . . . Lyrer, P. A. (2006). Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. *Stroke*, *37*(6), 1379-1384.
Farah, M. J., & McClelland, J. L. (1991). A computation model of semantic memory impairment - modality specificity and emergent category specificity. *Journal of Experimental Psychology-General*, *120*(4), 339-357.

Fareed, M., Suri, K., & Qureshi, A. (2012). Recruitment of ischemic stroke patients in clinical trials in general practice and implications for generalizability of results. *J Vasc Interv Neurol*, *5*(1), 27-32.

Fearnley, J. M., & Lees, A. J. (1991). Aging and Parkinsons Disease - substantia-nigra regional selectivity. *Brain*, *114*, 2283-2301.

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proc Natl Acad Sci U S A*, *108*(39), 16428-16433.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2012). Language-selective and domain-general regions lie side by side within Broca's area. *Curr Biol*, *22*(21), 2059-2062.

Fedorenko, E., Duncan, J., & Kanwisher, N. (2013). Broad domain generality in focal regions of frontal and parietal cortex. *Proc Natl Acad Sci U S A*, *110*(41), 16616-16621.

Ferbert, A., Priori, A., Rothwell, J. C., Day, B. L., Colebatch, J. G., & Marsden, C. D. (1992). Interhemispheric inhibition of the human motor cortex. *J Physiol*, *453*, 525-546.

Finger, S., Buckner, R. L., & Buckingham, H. (2003). Does the right hemisphere take over after damage to Broca's area? the Barlow case of 1877 and its history. *Brain and Language*, *85*(3), 385-395.

Fischer-Baum, S., Jang, A., & Kajander, D. (2017). The Cognitive Neuroplasticity of Reading Recovery following Chronic Stroke: A Representational Similarity Analysis Approach. *Neural Plasticity*, *2761913*.

Fridriksson, J. (2010). Preservation and modulation of specific left hemisphere regions is vital for treated recovery from anomia in stroke. *Journal of Neuroscience*, *30*(35), 11558-11564.

Fridriksson, J., Moser, D., Ryalls, J., Bonilha, L., Rorden, C., & Baylis, G. (2009). Modulation of frontal lobe speech areas associated with the production and perception of speech movements. *Journal of Speech, Language, and Hearing Research*, *52*(3), 812-819.

Fridriksson, J., Yourganov, G., Bonilha, L., Basilakos, A., Den Ouden, D. B., & Rorden, C. (2016). Revealing the dual streams of speech processing [Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(52), 15108-15113.

Friederici, A. D., & Gierhan, S. M. (2013). The language network. *Curr Opin Neurobiol*, *23*(2), 250-254.

Friston, K. J., Frith, C. D., Turner, R., & Frackowiak, R. S. J. (1995). Characterizing evoked hemodynamics with fMRI. *Neuroimage*, *2*(2), 157-165.

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen*, *141*(1), 2-18.

Fu, Y., Liu, Q., Anrather, J., & Shi, F. D. (2015). Immune interventions in stroke. *Nature Reviews Neurology*, *11*(9), 524-535.

Gajardo-Vidal, A., Lorca-Puls, D. L., Hope, T. M. H., Parker Jones, O., Seghier, M. L., Prejawa, S., . . . Price, C. J. (2018). How right hemisphere damage after stroke can impair speech comprehension. *Brain*, 141(12), 3389-3404.

Geranmayeh, F., Brownsett, S. L. E., & Wise, R. J. S. (2014). Task-induced brain activity in aphasic stroke patients: What is driving recovery? *Brain*, *137*(10), 2632-2648.

Geranmayeh, F., Chau, T. W., Wise, R. J. S., Leech, R., & Hampshire, A. (2017). Domain-general subregions of the medial prefrontal cortex contribute to recovery of language after stroke. *Brain*, *140*(7), 1947-1958.

Geranmayeh, F., Leech, R., & Wise, R. J. (2015). Semantic retrieval during overt picture description: Left anterior temporal or the parietal lobe? *Neuropsychologia*, *76*, 125-135.

Geranmayeh, F., Leech, R., & Wise, R. J. S. (2016). Network dysfunction predicts speech production after left hemisphere stroke. *Neurology*, *86*(14), 1296-1305.

Geranmayeh, F., Wise, R. J., Leech, R., & Murphy, K. (2015). Measuring vascular reactivity with

breath-holds after stroke: a method to aid interpretation of group-level BOLD signal changes in longitudinal fMRI studies. *Hum Brain Mapp*, *36*(5), 1755-1771.
Geranmayeh, F., Wise, R. J., Mehta, A., & Leech, R. (2014). Overlapping networks engaged during spoken language production and its cognitive control. *J Neurosci*, *34*(26), 8728-8740.
Geschwind, N. (1970). The organization of language and the brain. *Science*, *170*(3961), 940-944.
Gilmore, N., Meier, E. L., Johnson, J. P., & Kiran, S. (2019). Non-linguistic cognitive factors predict treatment-induced recovery in chronic post-stroke aphasia. *Arch Phys Med Rehabil*.
Gong, G. L., He, Y., Concha, L., Lebel, C., Gross, D. W., Evans, A. C., & Beaulieu, C. (2009). Mapping Anatomical Connectivity Patterns of Human Cerebral Cortex Using In Vivo Diffusion Tensor Imaging Tractography. *Cerebral Cortex*, *19*(3), 524-536.
Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed. ed.). Lea & Febiger.
Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychol Sci*, *13*(5), 425-430.
Gowers, W. R. (1893). A manual of diseases of the nervous system: vol. 2. In (2nd edition ed., pp. 110-125). P. Blakiston.
Gracco, V. L., Tremblay, P., & Pike, B. (2005). Imaging speech production using fMRI. *Neuroimage*, *26*(1), 294-301.
Grafman, J. (2000). Conceptualizing functional neuroplasticity. *Journal of Communication Disorders*, *33*(4), 345-356.
Grefkes, C., & Fink, G. R. (2011). Reorganization of cerebral networks after stroke: new insights from neuroimaging with connectivity approaches. *Brain*, 134(5), 1264-1276.
Griffis, J. C., Nenert, R., Allendorfer, J. B., Vannest, J., Holland, S., Dietz, A., & Szaflarski, J. P. (2017). The canonical semantic network supports residual language function in chronic post-stroke aphasia. *Human Brain Mapping*, *38*(3), 1636-1658.
Grönberg, A., Henriksson, I., & Lindgren, A. (2021). Accuracy of NIH Stroke Scale for diagnosing aphasia. *Acta Neurol Scand*, *143*(4), 375-382.
Hagoort, P., Wassenaar, M., & Brown, C. (2003). Real-time semantic compensation in patients with agrammatic comprehension: electrophysiological evidence for multiple-route plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7), 4340-4345.
Halai, A., De Dios Perez, B., Stefaniak, J., & Lambon Ralph, M. (2020). Comparing short and long batteries to assess deficits and their neural bases in stroke aphasia. *bioRxiv*. https://doi.org/10.1101/2020.11.24.395590
Halai, A., Woollams, A., & Lambon Ralph, M. (2020). Investigating the effect of changing parameters when building prediction models in post-stroke aphasia. *Nat Hum Behav*, 4(7), 725-735.
Halai, A. D., Parkes, L. M., & Welbourne, S. R. (2015). Dual-echo fMRI can detect activations in inferior temporal lobe during intelligible speech comprehension. *Neuroimage*, *122*, 214-221.
Halai, A. D., Woollams, A. M., & Lambon Ralph, M. A. (2017). Using principal component analysis to capture individual differences within a unified neuropsychological model of chronic post-stroke aphasia: Revealing the unique neural correlates of speech fluency, phonology and semantics. *Cortex*, *86*(275-289).
Hall, D. A., Haggard, M. P., Akeroyd, M. A., Palmer, A. R., Summerfield, A. Q., Elliott, M. R., . . . Bowtell, R. W. (1999). "Sparse" temporal sampling in auditory fMRI. *Hum Brain Mapp*, *7*(3), 213-223.

Hallam, G. P., Thompson, H. E., Hymers, M., Millman, R. E., Rodd, J. M., Lambon Ralph, M. A., . . . Jefferies, E. (2018). Task-based and resting-state fMRI reveal compensatory network changes following damage to left inferior frontal gyrus. *Cortex*, *99*(150-165).

Hartman, J. (1981). Measurement of early spontaneous recovery from aphasia with stroke. *Annals of Neurology*, *9*(1), 89-91.

Hartwigsen, G. (2018). Flexible Redistribution in Cognitive Networks. *Trends Cogn Sci*, *22*(8), 687-698.

Hartwigsen, G., Bzdok, D., Klein, M., Wawrzyniak, M., Stockert, A., Wrede, K., . . . Saur, D. (2017). Rapid short-term reorganization in the language network. *eLife*, *6*, 24.

Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425-2430.

Haynes, J. D. (2015). A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. *Neuron*, *87*(2), 257-270.

Haynes, J. D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Curr Biol*, *17*(4), 323-328.

He, B. J., Snyder, A. Z., Vincent, J. L., Epstein, A., Shulman, G. L., & Corbetta, M. (2007). Breakdown of functional connectivity in frontoparietal networks underlies behavioral deficits in spatial neglect. *Neuron*, *53*(6), 905-918.

Hebart, M. N., Görgen, K., & Haynes, J. D. (2014). The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Front Neuroinform*, *8*, 88.

Heikkinen, P. H., Pulvermuller, F., Makela, J. P., Ilmoniemi, R. J., Lioumis, P., Kujala, T., . . . Klippi, A. (2019). Combining rTMS with intensive language-action therapy in chronic aphasia: A randomized controlled trial. *Frontiers in Neuroscience*, 12, 1036.

Heilman, K. M. (2006). Aphasia and the diagram makers revisited: an update of information processing models. *J Clin Neurol*, *2*(3), 149-162.

Heiss, W. D., Hartmann, A., Rubi-Fessen, I., Anglade, C., Kracht, L., Kessler, J., . . . Thiel, A. (2013). Noninvasive brain stimulation for treatment of right- and left-handed poststroke aphasics. *Cerebrovasc Dis*, *36*(5-6), 363-372.

Heiss, W. D., Kessler, J., Thiel, A., Ghaemi, M., & Karbe, H. (1999). Differential capacity of left and right hemispheric areas for compensation of poststroke aphasia. *Annals of Neurology*, *45*(4), 430-438.

Heiss, W. D., & Thiel, A. (2006). A proposed regional hierarchy in recovery of post-stroke aphasia. *Brain and Language*, *98*(1), 118-123.

Herbet, G., Maheu, M., Costi, E., Lafargue, G., & Duffau, H. (2016). Mapping neuroplastic potential in brain-damaged patients. *Brain*, *139*(3), 829-844.

Hickok, G., Okada, K., Barr, W., Pa, J., Rogalsky, C., Donnelly, K., . . . Grant, A. (2008). Bilateral capacity for speech sound processing in auditory comprehension: evidence from Wada procedures. *Brain Lang*, *107*(3), 179-184.

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

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nat Rev Neurosci*, *8*(5), 393-402.

Hillis, A. E., Beh, Y. Y., Sebastian, R., Breining, B., Tippett, D. C., Wright, A., . . . Fridriksson, J. (2018). Predicting recovery in acute poststroke aphasia. *Annals of Neurology*, *83*(3), 612-622.

Hillis, A. E., & Heidler, J. (2002). Mechanisms of early aphasia recovery. *Aphasiology*, *16*(9), 885-895.

Hillis, A. E., Wityk, R. J., Barker, P. B., Beauchamp, N. J., Gailloud, P., Murphy, K., . . . Metter, E. J. (2002). Subcortical aphasia and neglect in acute stroke: the role of cortical hypoperfusion. *Brain, 125*(5), 1094-1104.

Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. *Brain, 115 (6),* 1783-1806.

Hodkinson, D. J., Bungert, A., Bowtell, R., Jackson, S. R., & Jung, J. (2021). Operculo-insular and anterior cingulate plasticity induced by transcranial magnetic stimulation in the human motor cortex: a dynamic casual modeling study. *J Neurophysiol, 125*(4), 1180-1190.

Hoffman, P., Binney, R. J., & Lambon Ralph, M. A. (2015). Differing contributions of inferior prefrontal and anterior temporal cortex to concrete and abstract conceptual knowledge. *Cortex, 63*, 250-266.

Hoffman, P., & Tamm, A. (2020). Barking up the right tree: Univariate and multivariate fMRI analyses of homonym comprehension. *Neuroimage, 219*, 117050.

Hon, N., Epstein, R. A., Owen, A. M., & Duncan, J. (2006). Frontoparietal activity with minimal decision and control. *J Neurosci, 26*(38), 9805-9809.

Hope, T. M. H., Friston, K., Price, C. J., Leff, A. P., Rotshtein, P., & Bowman, H. (2019). Recovery after stroke: not so proportional after all? *Brain, 142*, 15-22.

Hope, T. M. H., Leff, A. P., Prejawa, S., Bruce, R., Haigh, Z., Lim, L., . . . Price, C. J. (2017). Right hemisphere structural adaptation and changing language skills years after left hemisphere stroke. *Brain, 140*(6), 1718-1728.

Humphreys, G. F., & Gennari, S. P. (2014). Competitive mechanisms in sentence processing: common and distinct production and reading comprehension networks linked to the prefrontal cortex. *Neuroimage, 84*, 354-366.

Ito, K. L., Kim, H., & Liew, S. L. (2019). A comparison of automated lesion segmentation approaches for chronic stroke T1-weighted MRI data. *Hum Brain Mapp, 40*(16), 4669-4685.

Ivanova, M. V., Isaev, D. Y., Dragoy, O. V., Akinina, Y. S., Petrushevskiy, A. G., Fedina, O. N., . . . Dronkers, N. F. (2016). Diffusion-tensor imaging of major white matter tracts and their role in language processing in aphasia. *Cortex, 85*(165-181).

Jackson, J. B., Feredoes, E., Rich, A. N., Lindner, M., & Woolgar, A. (2021). Concurrent neuroimaging and neurostimulation reveals a causal role for dlPFC in coding of task-relevant information. *Commun Biol, 4*(1), 588.

Jackson, R. L. (2021). The neural correlates of semantic control revisited. *Neuroimage, 224*, 117444.

Jefferies, E., Patterson, K., Jones, R. W., & Lambon Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology, 23*(4), 492-499.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17*(2), 825-841.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. *Neuroimage, 62*(2), 782-790.

Jung, J., Bungert, A., Bowtell, R., & Jackson, S. R. (2020). Modulating Brain Networks With Transcranial Magnetic Stimulation Over the Primary Motor Cortex: A Concurrent TMS/fMRI Study. *Front Hum Neurosci, 14*, 31.

Jung, J., & Lambon Ralph, M. A. (2016). Mapping the Dynamic Network Interactions Underpinning Cognition: A cTBS-fMRI Study of the Flexible Adaptive Neural System for Semantics. *Cerebral Cortex, 26*(8), 3580-3590.

Kaiser, H. (1974). An index of factor simplicity. *Psychometrika, 39*, 31-36.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston naming test*. Lea & Febiger.
Kay, J., Lesser, R., & Coltheart, M. (1992). *PALPA: psycholinguistic assessments of language processing in aphasia.* Erlbaum.
Keidel, J. L., Welbourne, S. R., & Lambon Ralph, M. A. (2010). Solving the paradox of the equipotential and modular brain: a neurocomputational model of stroke vs. slow-growing glioma. *Neuropsychologia, 48*(6), 1716-1724.
Keuleers, E., & Brysbaert, M. (2010). Wuggy: a multilingual pseudoword generator. *Behav Res Methods, 42*(3), 627-633.
Kielar, A., Panamsky, L., Links, K. A., & Meltzer, J. A. (2015). Localization of electrophysiological responses to semantic and syntactic anomalies in language comprehension with MEG. *Neuroimage, 105*, 507-524.
Koechlin, E., & Jubault, T. (2006). Broca's area and the hierarchical organization of human behavior. *Neuron, 50*(6), 963-974.
Krainik, A., Hund-Georgiadis, M., Zysset, S., & von Cramon, D. Y. (2005). Regional impairment of cerebrovascular reactivity and BOLD signal in adults after stroke. *Stroke, 36*(6), 1146-1152.
Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proc Natl Acad Sci U S A, 103*(10), 3863-3868.
Kujala, T., Alho, K., Huotilainen, M., Ilmoniemi, R. J., Lehtokoski, A., Leinonen, A., . . . Naatanen, R. (1997). Electrophysiological evidence for cross-modal plasticity in humans with early- and late-onset blindness. *Psychophysiology, 34*(2), 213-216.
Kummerer, D., Hartwigsen, G., Kellmeyer, P., Glauche, V., Mader, I., Kloppel, S., . . . Saur, D. (2013). Damage to ventral and dorsal language pathways in acute aphasia. *Brain, 136*(2), 619-629.
Kundu, P., Brenowitz, N. D., Voon, V., Worbe, Y., Vértes, P. E., Inati, S. J., . . . Bullmore, E. T. (2013). Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proc Natl Acad Sci U S A, 110*(40), 16187-16192.
Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M., & Bandettini, P. A. (2012). Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *Neuroimage, 60*(3), 1759-1770.
Lacey, E. H., Skipper-Kallal, L. M., Xing, S., Fama, M. E., & Turkeltaub, P. E. (2017). Mapping Common Aphasia Assessments to Underlying Cognitive Processes and Their Neural Substrates. *Neurorehabilitation and Neural Repair, 31*(5), 442-450.
Lambon Ralph, M., Jefferies, E., Patterson, K., & Rogers, T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience, 18*, 42-55.
Lambon Ralph, M., Snell, C., Fillingham, J., Conroy, P., & Sage, K. (2010). Predicting the outcome of anomia therapy for people with aphasia post CVA: both language and cognitive status are key predictors. *Neuropsychol Rehabil, 20*(2), 289-305.
Lambon Ralph, M. A., McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R. (2001). No right to speak? The relationship between object naming and semantic impairment: neuropsychological evidence and a computational model. *J Cogn Neurosci, 13*(3), 341-356.
Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., . . . Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum Brain Mapp, 28*(11), 1194-1205.
Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., . . . Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Hum Brain Mapp, 10*(3), 120-131.
Laska, A. C., Hellblom, A., Murray, V., Kahan, T., & Von Arbin, M. (2001). Aphasia in acute

stroke and relation to outcome. *J Intern Med*, *249*(5), 413-422.
Lee, Y. S., Zreik, J. T., & Hamilton, R. H. (2017). Patterns of neural activity predict picture-naming performance of a patient with chronic aphasia. *Neuropsychologia*, *94*, 52-60.
Leff, A., Crinion, J., Scott, S., Turkheimer, F., Howard, D., & Wise, R. (2002). A physiological change in the homotopic cortex following left posterior temporal lobe infarction. *Ann Neurol*, *51*(5), 553-558.
Li, R., Perrachione, T. K., Tourville, J. A., & Kiran, S. (2021). Representation of semantic typicality in brain activation in healthy adults and individuals with aphasia: A multi-voxel pattern analysis. *Neuropsychologia*, *158*, 107893.
Li, T., Zeng, X., Lin, L., Xian, T., & Chen, Z. (2020). Effects of repetitive transcranial magnetic stimulation with different frequencies on post-stroke aphasia: A PRISMA-compliant meta-analysis. *Medicine (Baltimore)*, *99*(24), e20439.
Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J Neurosci Methods*, *264*, 47-56.
Lidzba, K., Schwilling, E., Grodd, W., Krägeloh-Mann, I., & Wilke, M. (2011). Language comprehension vs. language production: age effects on fMRI activation. *Brain Lang*, *119*(1), 6-15.
Lomas, J., & Kertesz, A. (1978). Patterns of spontaneous recovery in aphasic groups: a study of adult stroke patients. *Brain Lang*, *5*(3), 388-401.
Long, C., Sebastian, R., Faria, A. V., & Hillis, A. E. (2018). Longitudinal imaging of reading and naming recovery after stroke. *Aphasiology*, *32*(7), 839-854.
Lopez-Alonso, V., Cheeran, B., Rio-Rodriguez, D., & Fernandez-del-Olmo, M. (2014). Inter-individual Variability in Response to Non-invasive Brain Stimulation Paradigms. *Brain Stimulation*, *7*(3), 372-380.
Maas, M. B., Lev, M. H., Ay, H., Singhal, A. B., Greer, D. M., Smith, W. S., . . . Furie, K. L. (2012). The prognosis for aphasia in stroke. *Journal of Stroke and Cerebrovascular Diseases*, *21*(5), 350-357.
Manring, N. D., & Johnson, R. E. (1996). Modeling and designing a variable-displacement open-loop pump [Article]. *Journal of Dynamic Systems Measurement and Control-Transactions of the Asme*, *118*(2), 267-271.
Mansouri, F., Egner, T., & Buckley, M. (2017). Monitoring demands for executive control: shared functions between human and nonhuman primates. *Trends Neurosci*, *40*(1), 15-27.
Marebwa, B. K., Fridriksson, J., Yourganov, G., Feenaughty, L., Rorden, C., & Bonilha, L. (2017). Chronic post-stroke aphasia severity is determined by fragmentation of residual white matter networks. *Sci Rep*, *7*(1), 8188.
Margulies, D. S., Ghosh, S. S., Goulas, A., Falkiewicz, M., Huntenburg, J. M., Langs, G., . . . Smallwood, J. (2016). Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci U S A*, *113*(44), 12574-12579.
Marshall, J. F. (1984). Brain function: neural adaptations and recovery from injury. *Annu Rev Psychol*, *35*, 277-308.
Mattioli, F., Ambrosi, C., Mascaro, L., Scarpazza, C., Pasquali, P., Frugoni, M., . . . Gasparotti, R. (2014). Early aphasia rehabilitation is associated with functional reactivation of the left inferior frontal gyrus: a pilot study. *Stroke*, *45*(2), 545-552.
Mazoyer, B., Zago, L., Jobard, G., Crivello, F., Joliot, M., Perchey, G., . . . Tzourio-Mazoyer, N. (2014). Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PLoS One*, *9*(6), e101165.
McKeown, M. J., Hansen, L. K., & Sejnowsk, T. J. (2003). Independent component analysis of functional MRI: what is signal and what is noise? *Curr Opin Neurobiol*, *13*(5), 620-629.

Meier, E. L., Johnson, J. P., & Kiran, S. (2018). Left frontotemporal effective connectivity during semantic feature judgments in patients with chronic aphasia and age-matched healthy controls. *Cortex*, *108*(173-192).

Meier, E. L., Johnson, J. P., Pan, Y., & Kiran, S. (2019). A lesion and connectivity-based hierarchical model of chronic aphasia recovery dissociates patients and healthy controls. *NeuroImage: Clinical*, 23, 101919.

Meinzer, M., Flaisch, T., Seeds, L., Harnish, S., Antonenko, D., Witte, V., . . . Crosson, B. (2012). Same modulation but different starting points: Performance modulates age differences in inferior frontal cortex activity during word-retrieval. *PLoS ONE*, *7*(3), e33631.

Mementi, L., Gierhan, S., Segaert, K., & Hagoort, P. (2011). Shared language: overlap and segregation of the neuronal infrastructure for speaking and listening revealed by functional MRI. *Psychol Sci*, *22*(9), 1173-1182.

Mirman, D., Chen, Q., Zhang, Y., Wang, Z., Faseyitan, O. K., Coslett, H. B., & Schwartz, M. F. (2015). Neural organization of spoken language revealed by lesion-symptom mapping. *Nature communications*, *6*, 6762.

Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage*, *53*(1), 103-118.

Mitchell, R. L., Vidaki, K., & Lavidor, M. (2016). The role of left and right dorsolateral prefrontal cortex in semantic processing: A transcranial direct current stimulation study. *Neuropsychologia*, *91*, 480-489.

Molad, J., Kliper, E., Korczyn, A. D., Ben Assayag, E., Ben Bashat, D., Shenhar-Tsarfaty, S., . . . Auriel, E. (2017). Only White Matter Hyperintensities Predicts Post-Stroke Cognitive Performances Among Cerebral Small Vessel Disease Markers: Results from the TABASCO Study. *Journal of Alzheimers Disease*, *56*(4), 1293-1299.

Morcom, A. M., & Henson, R. N. A. (2018). Increased Prefrontal Activity with Aging Reflects Nonspecific Neural Responses Rather than Compensation. *J Neurosci*, *38*(33), 7303-7313.

Mumford, J. A. (2017). A comprehensive review of group level model performance in the presence of heteroscedasticity: Can a single model control Type I errors in the presence of outliers? *Neuroimage*, *147*, 658-668.

Murphy, T. H., & Corbett, D. (2009). Plasticity during stroke recovery: from synapse to behaviour. *Nat Rev Neurosci*, *10*(12), 861-872.

Murray, L. L. (2000). The effects of varying attentional demands on the word retrieval skills of adults with aphasia, right hemisphere brain damage, or no brain damage. *Brain Lang*, *72*(1), 40-72.

Murray, L. L. (2012). Attention and other cognitive deficits in aphasia: Presence and relation to language and communication measures. *American Journal of Speech-Language Pathology*, *21*(2), s51-s64.

Müller, V. I., Cieslik, E. C., Laird, A. R., Fox, P. T., Radua, J., Mataix-Cols, D., . . . Eickhoff, S. B. (2018). Ten simple rules for neuroimaging meta-analysis. *Neurosci Biobehav Rev*, *84*, 151-161.

Naidech, A. M., Kreiter, K. T., Janjua, N., Ostapkovich, N., Parra, A., Commichau, C., . . . Fitzsimmons, B. F. M. (2005). Phenytoin exposure is associated with functional and cognitive disability after subarachnoid hemorrhage. *Stroke*, *36*(3), 583-587.

Nenert, R., Allendorfer, J. B., Martin, A. M., Banks, C., Vannest, J., Holland, S. K., . . . Szaflarski, J. P. (2018). Longitudinal fMRI study of language recovery after a left hemispheric ischemic stroke. *Restorative Neurology and Neuroscience*, *36*(3), 359-385.

Niven, J. E. (2016). Neuronal energy consumption: biophysics, efficiency and evolution. *Current*

Opinion in Neurobiology, *41*, 129-135.
Nyberg, L., Andersson, M., Kauppi, K., Lundquist, A., Persson, J., Pudas, S., & Nilsson, L. G. (2014). Age-related and Genetic Modulation of Frontal Cortex Efficiency. *Journal of Cognitive Neuroscience*, *26*(4), 746-754.
O'Brien, P., Sellar, R. J., & Wardlaw, J. M. (2004). Fogging on T2-weighted MR after acute ischaemic stroke: how often might this occur and what are the implications? *Neuroradiology*, *46*(8), 635-641.
Olszowy, W., Aston, J., Rua, C., & Williams, G. B. (2019). Accurate autocorrelation modeling substantially improves fMRI reliability. *Nat Commun*, *10*(1), 1220.
Patterson, K., & Lambon Ralph, M. A. (1999). Selective disorders of reading? *Curr Opin Neurobiol*, *9*(2), 235-239.
Pedersen, P. M., Jørgensen, H. S., Nakayama, H., Raaschou, H. O., & Olsen, T. S. (1995). Aphasia in acute stroke: incidence, determinants, and recovery. *Ann Neurol*, *38*(4), 659-666.
Perrachione, T. K., & Ghosh, S. S. (2013). Optimized design and analysis of sparse-sampling fMRI experiments. *Frontiers in Neuroscience*, *7*, 55.
Phan, T., Donnan, G., Wright, P., & Reutens, D. (2005). A digital map of middle cerebral artery infarcts associated with middle cerebral artery trunk and branch occlusion. *Stroke*, *36*(5), 986-991.
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychol Rev*, *103*(1), 56-115.
Pobric, G., Jefferies, E., & Ralph, M. A. (2007). Anterior temporal lobes mediate semantic representation: mimicking semantic dementia by using rTMS in normal participants. *Proc Natl Acad Sci U S A*, *104*(50), 20137-20141.
Postman-Caucheteux, W. A., Birn, R. M., Pursley, R. H., Butman, J. A., Solomon, J. M., Picchioni, D., . . . Braun, A. R. (2010). Single-trial fMRI shows contralesional activity linked to overt naming errors in chronic aphasic patients. *Journal of Cognitive Neuroscience*, *22*(6), 1299-1318.
Price, C. J., & Friston, K. J. (2002). Degeneracy and cognitive anatomy. *Trends Cogn Sci*, *6*(10), 416-421.
Price, C. J., Warburton, E. A., Moore, C. J., Frackowiak, R. S., & Friston, K. J. (2001). Dynamic diaschisis: anatomically remote and context-sensitive human brain lesions. *J Cogn Neurosci*, *13*(4), 419-429.
Pritchett, B. L., Hoeflin, C., Koldewyn, K., Dechter, E., & Fedorenko, E. (2018). High-level language processing regions are not engaged in action observation or imitation. *J Neurophysiol*, *120*(5), 2555-2570.
Puente, A. N., Lindbergh, C. A., & Miller, L. S. (2015). The Relationship Between Cognitive Reserve and Functional Ability is Mediated by Executive Functioning in Older Adults. *Clinical Neuropsychologist*, *29*(1), 67-81.
Pustina, D., Coslett, H. B., Turkeltaub, P. E., Tustison, N., Schwartz, M. F., & Avants, B. (2016). Automated segmentation of chronic stroke lesions using LINDA: Lesion identification with neighborhood data analysis. *Human Brain Mapping*, *37*(4), 1405-1421.
Qiu, W. H., Wu, H. X., Yang, Q. L., Kang, Z., Chen, Z. C., Li, K., . . . Chen, S. Q. (2017). Evidence of cortical reorganization of language networks after stroke with subacute Broca's aphasia: A blood oxygenation level dependent-functional magnetic resonance imaging study. *Neural Regeneration Research*, *12*(1), 109-117.
Radman, N., Mouthon, M., Di Pietro, M., Gaytanidis, C., Leemann, B., Abutalebi, J., & Annoni, J. M. (2016). The Role of the Cognitive Control System in Recovery from Bilingual

Aphasia: A Multiple Single-Case fMRI Study. *Neural Plasticity, 8797086*.
Rafiei, F., Safrin, M., Wokke, M. E., Lau, H., & Rahnev, D. (2021). Transcranial magnetic stimulation alters multivoxel patterns in the absence of overall activity changes. *Hum Brain Mapp*, 42(12), 3804-3820.
Raichle, M. E., & Gusnard, D. A. (2002). Appraising the brain's energy budget. *Proceedings of the National Academy of Sciences of the United States of America, 99*(16), 10237-10239.
Rajtar-Zembaty, A., Przewoznik, D., Bober-Plonka, B., Starowicz-Filip, A., Rajtar-Zembaty, J., Nowak, R., & Przewlocki, R. (2015). Application of the Trail Making Test in the assessment of cognitive flexibility in patients with speech disorders after ischaemic cerebral stroke. *Aktualnosci Neurologiczne, 15*(1), 11-17.
Ramanathan, D., Tuszynski, M. H., & Conner, J. M. (2009). The Basal Forebrain Cholinergic System Is Required Specifically for Behaviorally Mediated Cortical Map Plasticity. *Journal of Neuroscience, 29*(18), 5992-6000.
Ramsey, L. E., Siegel, J. S., Lang, C. E., Strube, M., Shulman, G. L., & Corbetta, M. (2017). Behavioural clusters and predictors of performance during recovery from stroke. *Nat Hum Behav*, 1, 0038.
Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci, 12*(6), 718-724.
Raven, J. C. (1962). *Advanced progressive matrices, set II*. H. K. Lewis.
Ren, C. L., Zhang, G. F., Xia, N., Jin, C. H., Zhang, X. H., Hao, J. F., . . . Cai, D. L. (2014). Effect of low-frequency rTMS on aphasia in stroke patients: a meta-analysis of randomized controlled trials. *PLoS One, 9*(7), e102557.
Rice, G. E., Caswell, H., Moore, P., Lambon Ralph, M. A., & Hoffman, P. (2018). Revealing the Dynamic Modulations That Underpin a Resilient Neural Network for Semantic Cognition: An fMRI Investigation in Patients With Anterior Temporal Lobe Resection. *Cereb Cortex, 28*(8), 3004-3016.
Rice, G. E., Lambon Ralph, M. A., & Hoffman, P. (2015). The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies. *Cereb Cortex, 25*(11), 4374-4391.
Robson, H., Zahn, R., Keidel, J. L., Binney, R. J., Sage, K., & Lambon Ralph, M. A. (2014). The anterior temporal lobes support residual comprehension in Wernicke's aphasia. *Brain, 137*(3), 931-943.
Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language, 72*(3), 193-218.
Rosen, H. J., Petersen, S. E., Linenweber, M. R., Snyder, A. Z., White, D. A., Chapman, L., . . . Corbetta, M. D. (2000). Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology, 55*(12), 1883-1894.
Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production - procedure and data. *Brain and Language, 37*(3), 440-479.
Sajid, N., Parr, T., Hope, T., Price, C., & Friston, K. (2020). Degeneracy and redundancy in active inference. *Cereb Cortex*, 30(11), 5750-5766.
Salvalaggio, A., De Filippo De Grazia, M., Zorzi, M., Thiebaut de Schotten, M., & Corbetta, M. (2020). Post-stroke deficit prediction from lesion and indirect structural and functional disconnection. *Brain, 143*(7), 2173-2188.
Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., . . . Weiller, C. (2008). Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A, 105*(46), 18035-18040.
Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller,

C. (2006). Dynamics of language reorganization after stroke. *Brain*, *129*(6), 1371-1384.
Saur, D., Ronneberger, O., Kummerer, D., Mader, I., Weiller, C., & Kloppel, S. (2010). Early functional magnetic resonance imaging activations predict language outcome after stroke. *Brain*, *133*(4), 1252-1264.
Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., . . . Weiller, C. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *Neuroimage*, *49*(4), 3187-3197.
Schapiro, A. C., McClelland, J. L., Welbourne, S. R., Rogers, T. T., & Lambon Ralph, M. A. (2013). Why bilateral damage is worse than unilateral damage to the brain. *J Cogn Neurosci*, *25*(12), 2107-2123.
Schofield, T. M., Penny, W. D., Stephan, K. E., Crinion, J. T., Thompson, A. J., Price, C. J., & Leff, A. P. (2012). Changes in auditory feedback connections determine the severity of speech processing deficits after stroke. *Journal of Neuroscience*, *32*(12), 4260-4270.
Schumacher, R., Halai, A. D., & Lambon Ralph, M. A. (2019). Assessing and mapping language, attention and executive multidimensional deficits in stroke aphasia. *Brain*, *142*(10), 3202-3216.
Schwartz, M. F., Faseyitan, O., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain*, *135*(Pt 12), 3799-3814.
Seghier, M. L., Patel, E., Prejawa, S., Ramsden, S., Selmer, A., Lim, L., . . . Price, C. J. (2016). The PLORAS Database: A data repository for Predicting Language Outcome and Recovery After Stroke. *Neuroimage*, *124*(B), 1208-1212.
Seghier, M. L., Ramlackhansingh, A., Crinion, J., Leff, A. P., & Price, C. J. (2008). Lesion identification using unified segmentation-normalisation models and fuzzy clustering. *Neuroimage*, *41*(4), 1253-1266.
Seniow, J., Waldowski, K., Lesniak, M., Iwanski, S., Czepiel, W., & Czlonkowska, A. (2013). Transcranial magnetic stimulation combined with speech and language training in early aphasia rehabilitation: A randomized double-blind controlled pilot study. *Topics in Stroke Rehabilitation*, *20*(3), 250-261.
Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge University Press.
Sharp, D. J., Turkheimer, F. E., Bose, S. K., Scott, S. K., & Wise, R. J. (2010). Increased frontoparietal integration after stroke and cognitive recovery. *Annals of Neurology*, *68*(5), 753-756.
Siegel, J. S., Ramsey, L. E., Snyder, A. Z., Metcalf, N. V., Chacko, R. V., Weinberger, K., . . . Corbetta, M. (2016). Disruptions of network connectivity predict impairment in multiple behavioral domains after stroke. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(30), E4367-4376.
Siegel, J. S., Seitzman, B. A., Ramsey, L. E., Ortega, M., Gordon, E. M., Dosenbach, N. U. F., . . . Corbetta, M. (2018). Re-emergence of modular brain networks in stroke recovery. *Cortex*, 101, 44-59.
Simmonds, A. J., Leech, R., Collins, C., Redjep, O., & Wise, R. J. (2014). Sensory-motor integration during speech production localizes to both left and right plana temporale. *J Neurosci*, *34*(39), 12963-12972.
Skipper-Kallal, L. M., Lacey, E. H., Xing, S., & Turkeltaub, P. E. (2017a). Functional activation independently contributes to naming ability and relates to lesion site in post-stroke aphasia. *Human Brain Mapping*, *38*(4), 2051-2066.
Skipper-Kallal, L. M., Lacey, E. H., Xing, S., & Turkeltaub, P. E. (2017b). Right Hemisphere Remapping of Naming Functions Depends on Lesion Size and Location in Poststroke Aphasia. *Neural Plasticity*, *8740353*.
Sliwinska, M. W., Violante, I. R., Wise, R. J. S., Leech, R., Devlin, J. T., Geranmayeh, F., &

Hampshire, A. (2017). Stimulating Multiple-Demand Cortex Enhances Vocabulary Learning. *Journal of Neuroscience*, *37*(32), 7606-7618.
Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143-155.
Southwell, D. G., Hervey-Jumper, S. L., Perry, D. W., & Berger, M. S. (2016). Intraoperative mapping during repeat awake craniotomy reveals the functional plasticity of adult cortex. *J Neurosurg*, *124*(5), 1460-1469.
SSNAP. (2017). *Sentinel Stroke National Audit Programme (SSNAP)*. Retrieved 03/07/2018 from https://www.strokeaudit.org/results/Clinical-audit/National-Results.aspx
Stefaniak, J. D., Alyahya, R. S. W., & Lambon Ralph, M. A. (2021). Language networks in aphasia and health: A 1000 participant Activation Likelihood Estimation meta-analysis. *NeuroImage*, *117960*.
Stefaniak, J. D., Halai, A. D., & Lambon Ralph, M. A. (2020). The neural and neurocomputational bases of recovery from post-stroke aphasia. *Nat Rev Neurol*, *16*(1), 43-55.
Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease [Article]. *Lancet Neurology*, *11*(11), 1006-1012.
Stockert, A., Wawrzyniak, M., Klingbeil, J., Wrede, K., Kummerer, D., Hartwigsen, G., . . . Saur, D. (2020). Dynamics of language reorganization after left temporo-parietal and frontal stroke. *Brain*, *143*(3), 844-861.
Su, C.-Y., Wuang, Y.-P., Lin, Y.-H., & Su, J.-H. (2015). The role of processing speed in post-stroke cognitive dysfunction. *Archives of Clinical Neuropsychology*, *30*(2), 148-160.
Swinburn, K., Baker, G., & Howard, D. (2005). *CAT: the comprehensive aphasia test*. Psychology Press.
Szaflarski, J. P., Allendorfer, J. B., Banks, C., Vannest, J., & Holland, S. K. (2013). Recovered vs. not-recovered from post-stroke aphasia: The contributions from the dominant and non-dominant hemispheres. *Restorative Neurology and Neuroscience*, *31*(4), 347-360.
Szaflarski, J. P., Eaton, K., Ball, A. L., Banks, C., Vannest, J., Allendorfer, J. B., . . . Holland, S. K. (2011). Poststroke aphasia recovery assessed with functional magnetic resonance imaging and a picture identification task. *Journal of Stroke and Cerebrovascular Diseases*, *20*(4), 336-345.
Szalay, G., Martinecz, B., Lénárt, N., Környei, Z., Orsolits, B., Judák, L., . . . Dénes, Á. (2016). Microglia protect against brain injury and their selective elimination dysregulates neuronal network activity after stroke. *Nat Commun*, *7*, 11499.
Tao, Y., & Rapp, B. (2019). The effects of lesion and treatment-related recovery on functional network modularity in post-stroke dysgraphia. *NeuroImage: Clinical*, 23, 101865.
Teufel, C., & Fletcher, P. C. (2016). The promises and pitfalls of applying computational models to neurological and psychiatric disorders. *Brain*, *139*(10), 2600-2608.
Thiel, A., Habedank, B., Herholz, K., Kessler, J., Winhuisen, L., Haupt, W. F., & Heiss, W. D. (2006). From the left to the right: How the brain compensates progressive loss of language function. *Brain Lang*, *98*(1), 57-65.
Thiel, A., Hartmann, A., Rubi-Fessen, I., Anglade, C., Kracht, L., Weiduschat, N., . . . Heiss, W. D. (2013). Effects of noninvasive brain stimulation on language networks and recovery in early poststroke aphasia. *Stroke*, *44*(8), 2240-2246.
Thiel, A., Herholz, K., Koyuncu, A., Ghaemi, M., Kracht, L. W., Habedank, B., & Heiss, W. D. (2001). Plasticity of language networks in patients with brain tumors: a positron emission tomography activation study. *Ann Neurol*, *50*(5), 620-629.
Thomalla, G., Boutitie, F., Fiebach, J. B., Simonsen, C. Z., Nighoghossian, N., Pedraza, S., . . . Gerloff, C. (2017). Effect of informed consent on patient characteristics in a stroke thrombolysis trial. *Neurology*, *89*(13), 1400-1407.

Tononi, G., Sporns, O., & Edelman, G. M. (1999). Measures of degeneracy and redundancy in biological networks. *Proc Natl Acad Sci U S A*, *96*(6), 3257-3262.

Tsouli, S., Kyritsis, A. P., Tsagalis, G., Virvidaki, E., & Vemmos, K. N. (2009). Significance of aphasia after first-ever acute stroke: impact on early and late outcomes. *Neuroepidemiology*, *33*(2), 96-102.

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in Activation Likelihood Estimation meta-analyses. *Hum Brain Mapp*, *33*(1), 1-13.

Turkeltaub, P. E., Messing, S., Norise, C., & Hamilton, R. H. (2011). Are networks for residual language function and recovery consistent across aphasic patients? *Neurology*, *76*(20), 1726-1734.

Tyler, L. K., Marslen-Wilson, W., & Stamatakis, E. A. (2005). Dissociating neuro-cognitive component processes: voxel-based correlational methodology. *Neuropsychologia*, *43*(5), 771-778.

Tyler, L. K., Wright, P., Randall, B., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Reorganization of syntactic processing following left-hemisphere brain damage: does right-hemisphere activity preserve function? *Brain*, *133*(11), 3396-3408.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., . . . Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, *15*(1), 273-289.

Ueno, T., Saito, S., Rogers, T., & Lambon Ralph, M. (2011). Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, *72*(2), 385-396.

Uiterwijk, R., van Oostenbrugge, R. J., Huijts, M., De Leeuw, P. W., Kroon, A. A., & Staals, J. (2016). Total Cerebral Small Vessel Disease MRI Score Is Associated with Cognitive Decline in Executive Function in Patients with Hypertension. *Frontiers in Aging Neuroscience*, 8, 301.

Umarova, R., & Thomalla, G. (2020). Indirect connectome-based prediction of post-stroke deficits: prospects and limitations. *Brain*, *143*(7), 1966-1970.

van Oers, C., van der Worp, H. B., Kappelle, L. J., Raemaekers, M. A. H., Otte, W. M., & Dijkhuizen, R. M. (2018). Etiology of language network changes during recovery of aphasia after stroke. *Scientific Reports*, *8*(1), 856.

van Oers, C. A., Vink, M., van Zandvoort, M. J., van der Worp, H. B., de Haan, E. H., Kappelle, L. J., . . . Dijkhuizen, R. M. (2010). Contribution of the left and right inferior frontal gyrus in recovery from aphasia. A functional MRI study in stroke patients with preserved hemodynamic responsiveness. *Neuroimage*, *49*(1), 885-893.

van Oers, C. A. M. M., Vink, M., van Zandvoort, M. J. E., van der Worp, H. B., de Haan, E. H. F., Kappelle, L. J., . . . Dijkhuizen, R. M. (2010). Contribution of the left and right inferior frontal gyrus in recovery from aphasia. A functional MRI study in stroke patients with preserved hemodynamic responsiveness. *NeuroImage*, *49*(1), 885-893.

Wade, D. T., Hewer, R. L., David, R. M., & Enderby, P. M. (1986). Aphasia after stroke: natural history and associated deficits. *J Neurol Neurosurg Psychiatry*, *49*(1), 11-16.

Wager, T. D., Keller, M. C., Lacey, S. C., & Jonides, J. (2005). Increased sensitivity in neuroimaging analyses using robust regression. *Neuroimage*, *26*(1), 99-113.

Waldowski, K., Seniow, J., Lesniak, M., Iwanski, S., & Czlonkowska, A. (2012). Effect of low-frequency repetitive transcranial magnetic stimulation on naming abilities in early-stroke aphasic patients: A prospective, randomized, double-blind sham-controlled study. *The Scientific World Journal*, *518568*.

Walenski, M., Europa, E., Caplan, D., & Thompson, C. K. (2019). Neural networks for sentence comprehension and production: An ALE-based meta-analysis of neuroimaging studies. *Hum Brain Mapp*, 40(8), 2275-2304.

Warburton, E., Price, C. J., Swinburn, K., & Wise, R. J. (1999). Mechanisms of recovery from aphasia: evidence from positron emission tomography studies. *Journal of Neurology, Neurosurgery & Psychiatry*, *66*(2), 155-161.

Wechsler, D. A. (1987). *Wechsler memory scale—revised manual*. Psychological Corporation.

Weiller, C., Bormann, T., Saur, D., Musso, M., & Rijntjes, M. (2011). How the ventral pathway got lost: and what its recovery might mean. *Brain Lang*, *118*(1-2), 29-39.

Welbourne, S. R., Woollams, A. M., Crisp, J., & Ralph, M. A. L. (2011). The role of plasticity-related functional reorganization in the explanation of central dyslexias. *Cognitive Neuropsychology*, *28*(2), 65-108.

Wernicke, C. (1874). *Der aphasische Symptomencomplex, eine psychologische Studie auf anatomischer Basis*. Breslau: Cohn and Weigert.

Wieshmann, U. C., Clark, C. A., Symms, M. R., Franconi, F., Barker, G. J., & Shorvon, S. D. (1999). Reduced anisotropy of water diffusion in structural cerebral abnormalities demonstrated with diffusion tensor imaging. *Magnetic Resonance Imaging*, *17*(9), 1269-1274.

Wiethoff, S., Hamada, M., & Rothwell, J. C. (2014). Variability in Response to Transcranial Direct Current Stimulation of the Motor Cortex. *Brain Stimulation*, *7*(3), 468-475.

Wilson, S. M., Yen, M., & Eriksson, D. K. (2018). An adaptive semantic matching paradigm for reliable and valid language mapping in individuals with aphasia. *Human Brain Mapping*, *39*(8), 3285-3307.

Winhuisen, L., Thiel, A., Schumacher, B., Kessler, J., Rudolf, J., Haupt, W. F., & Heiss, W. D. (2005). Role of the contralateral inferior frontal gyrus in recovery of language function in poststroke aphasia: a combined repetitive transcranial magnetic stimulation and positron emission tomography study. *Stroke*, *36*(8), 1759-1763.

Woo, C. W., Krishnan, A., & Wager, T. D. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage*, *91*, 412-419.

Woodhead, Z. V., Crinion, J., Teki, S., Penny, W., Price, C. J., & Leff, A. P. (2017). Auditory training changes temporal lobe connectivity in 'Wernicke's aphasia': a randomised trial. *Journal of Neurology, Neurosurgery & Psychiatry*, *88*(7), 586-594.

Woodhead, Z. V. J., Kerry, S. J., Aguilar, O. M., Ong, Y. H., Hogan, J. S., Pappa, K., . . . Crinion, J. T. (2018). Randomized trial of iReadMore word reading training and brain stimulation in central alexia. *Brain*, *141*, 2127-2141.

Woolgar, A., Bor, D., & Duncan, J. (2013). Global Increase in Task-related Fronto-parietal Activity after Focal Frontal Lobe Lesion. *Journal of Cognitive Neuroscience*, *25*(9), 1542-1552.

Woollams, A. M., Madrid, G., & Lambon Ralph, M. A. (2017). Using neurostimulation to understand the impact of pre-morbid individual differences on post-lesion outcomes. *Proc Natl Acad Sci U S A*, *114*(46), 12279-12284.

Xing, S., Lacey, E. H., Skipper-Kallal, L. M., Jiang, X., Harris-Love, M. L., Zeng, J., & Turkeltaub, P. E. (2016). Right hemisphere grey matter structure and language outcomes in chronic left hemisphere stroke. *Brain*, *139*(1), 227-241.

Xing, S., Lacey, E. H., Skipper-Kallal, L. M., Zeng, J., & Turkeltaub, P. E. (2017). White matter correlates of auditory comprehension outcomes in chronic post-stroke aphasia. *Frontiers in Neurology*, *8*, 54.

Xu, J. S., Zhang, S., Calhoun, V. D., Monterosso, J., Li, C. S. R., Worhunsky, P. D., . . . Potenza, M. N. (2013). Task-related concurrent but opposite modulations of overlapping functional

networks as revealed by spatial IA. *Neuroimage*, *79*, 62-71.
Yagata, S. A., Yen, M., McCarron, A., Bautista, A., Lamair-Orosco, G., & Wilson, S. M. (2017). Rapid recovery from aphasia after infarction of Wernicke's area. *Aphasiology*, *31*(8), 951-980.
Yamamoto, A. K., Jones, O. P., Hope, T. M. H., Prejawa, S., Oberhuber, M., Ludersdorfer, P., . . . Price, C. J. (2019). A special role for the right posterior superior temporal sulcus during speech production. *Neuroimage*, 203, 116184.
Yao, L., Zhao, H., Shen, C., Liu, F., Qiu, L., & Fu, L. (2020). Low-Frequency Repetitive Transcranial Magnetic Stimulation in Patients With Poststroke Aphasia: Systematic Review and Meta-Analysis of Its Effect Upon Communication. *J Speech Lang Hear Res*, *63*(11), 3801-3815.
Zhang, J., Zhong, D., Xiao, X., Yuan, L., Li, Y., Zheng, Y., . . . Jin, R. (2021). Effects of repetitive transcranial magnetic stimulation (rTMS) on aphasia in stroke patients: A systematic review and meta-analysis. *Clin Rehabil*, *35*(8), 1103-1116.
Zhao, Y., Halai, A., & Lambon Ralph, M. (2020). Evaluating the granularity and statistical structure of lesions and behaviour in post-stroke aphasia. *Brain Comm*, 2(2), fcaa062.
Zumbansen, A., Peretz, I., & Hebert, S. (2014). Melodic intonation therapy: back to basics for future research. *Frontiers in Neurology*, *5*, 7.