NEURAL MECHANISMS UNDERLYING WORD LEARNING IN HEALTHY AGEING AND POST-STROKE APHASIA

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

2022

KATHERINE R. GORE

DIVISION OF NEUROSCIENCE AND EXPERIMENTAL PSYCHOLOGY

TABLE OF CONTENTS

Word count: 51,136

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The overarching aim of this thesis was to explore the neural underpinnings of word learning in neurotypical populations and post-stroke aphasia. Word learning has been previously proposed to fit within a powerful framework, the Complementary Learning Systems (CLS) model. This model considers the acquisition, consolidation, and generalisation of new knowledge, whereby a hippocampal system complements neocortical memory systems in a computational trade-off between learning episode specifics and regularities across episodes.

In this thesis, this proposed neural division of labour was tested in adults through an investigation of the consolidation and long-term retention of newly learned, native vocabulary and neuroimaging. Consistent with CLS theory, naming of newly learned items was functionally supported by a combination of regions associated with episodic memory and the language-semantic areas that supported established vocabulary (Chapter 1). Additionally, the results suggest that interactivity between the complementary episodic and domain-specific neocortical memory systems supporting new knowledge is time-limited (Chapter 2). The structural correlates of word learning were also explored. Associations between right-sided structural intensities and better behavioural performance were moderated by structural integrity of dominant left hemispheric language regions. These results are considered within a set of ideas termed "variable neuro-displacement", a principle that may be applicable across several different populations (Chapter 3). In post-stroke aphasia, the functional correlates of word re-learning were largely normal up to a critical point of damage to dominant language processing regions (Chapter 4). Similarly, novel, native word learning in aphasia largely mirrored that of healthy controls. Although the functional correlates were largely normal, structural abnormalities in the dorsal and ventral white matter pathways were associated with the longevity of learning retention (Chapter 5).

Overall, the results of the thesis provide evidence of a general framework of word learning in healthy adults and post-stroke aphasia, which fits within the CLS model of knowledge acquisition. This framework held up until a critical point of damage in post-stroke aphasia. The theoretical and clinical implications of these results are discussed.

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

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance Presentation of Theses Policy You are required to submit your thesis electronically Page 11 of 25 with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=2442 0), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.

ACKNOWLEDGEMENTS

I am very grateful to my supervisors Prof Matt Lambon Ralph and Dr Anna Woollams for their continuous and invaluable support, insight, and advice throughout this process. I have also gratefully benefitted from discussions with Dr Steffie Bruehl, Dr Ajay Halai and Dr Grace Rice.

I would like to express my gratitude the participants who gave up their time and energy for this thesis. Many generously spent time doing the exact things they find most difficult. Others enthusiastically humoured me, learning weird and wonderful words without complaint.

Further thanks must go to my PhD colleagues past and present, for always accepting my socialisation needs and office cartoons. I am particularly indebted to Josie Kearney and Jayesha Chudasama for their unwavering friendship and support every step of the way. Additionally, I am grateful to have shared this experience with Dr Chris Heath. A heartfelt thank you to my parents and sister for always believing I could do whatever I set my mind to, no matter how often my mind fluttered. And last, but by no means least, Dr Becky Jackson; I will never be able to thank you enough for being you.

This work was supported by a Medical Research Council National Productivity Investment Fund PhD studentship, with support from Bioxydyn Limited.

GENERAL INTRODUCTION

THESIS OVERVIEW

This thesis has been submitted in alternative thesis format. As such, Chapters 1 to 5 are written in the style of journal articles. Accordingly, each empirical chapter begins with the literature and issues specific to that chapter, followed by the chapter-specific research aims, Methods, Results, and Discussion. Thus, there is inevitable repetition to ensure that each chapter can stand alone as a journal article. Chapter 1 involved an fMRI study of novel, native word learning in healthy adults. Chapter 2 explored functional connectivity changes during consolidation of this older adult vocabulary acquisition. Chapter 3 investigates the underlying structural underpinnings of this word learning in older adults. In Chapter 4, the functional neural correlates of word re-learning in patients with post-stroke aphasia was investigated. Chapter 5 explored the functional and white matter connectivity correlates of novel word learning in mild post-stroke aphasia in a case series. A general discussion of how the aims of this thesis were addressed and areas of further exploration are raised in Chapter 6: General Discussion.

To provide some theoretical grounding of relevant psycholinguistic and neurocognitive language processes, the first two parts of this Introduction provide a brief general introduction to post-stroke aphasia and theories of speech production. The next two sections detail the relevant neuroscientific and aphasiological literature of two broad cross-cutting themes in this thesis: neurotypical vocabulary acquisition and vocabulary re-acquisition in patients with post-stroke aphasia (PWA). The final section details the key aims of this thesis.

POST-STROKE APHASIA

Traumatic brain injury, neurodegeneration and cerebrovascular stroke can result in acquired language impairments ("aphasia"). Approximately one-third of strokes result in aphasia (Berthier 2005). Language impairments in aphasia lead to communication difficulties, which have catastrophic effects on quality of life and social participation (Code 2003). Thus, speech and language therapy (SLT) often focuses on re-learning or learning vocabulary to increase the ability to retrieve words in everyday life and improve communication (Best et al. 2008). From both cognitive and clinical neuroscience perspectives, it is fundamentally important to understand both the cognitive and neural bases of vocabulary acquisition.

Patients with post-stroke aphasia (PWA) form a heterogeneous group with vast variability of language deficits, lesion volume and location. Category-based classification schemes have been developed to assist with clinical stratification of diagnosis, intervention-based management, and language research activities. The classical typology of aphasia arose from the standard neurological model of language, the Broca-Wernicke-Lichtheim-Geschwind model (Broca 1861; Wernicke 1874; Lichtheim 1885; Geschwind 1965). In this historical model, speech production arises from inferior frontal regions (Broca, 1861), whereas comprehension is supported by posterior temporal regions (Wernicke, 1874). White matter connectivity via the arcuate fasciculus was suggested to support the transfer of information between these regions (Geschwind 1965).

Thus, traditional models of aphasia stratified patients by localisation-based classifications. However, damage caused by stroke is often widespread and not focal to one region. Thus, the exact location and function of Broca's and Wernicke's areas are disputed (Dronkers et al. 2007; Mesulam et al. 2015). Therefore, patients are now classified based upon their language deficits at clinical presentation through a broad array of neuropsychological testing. The Boston Diagnostic Aphasia Examination (BDAE; Kaplan et al., 1983) classifies patients into seven aphasic syndromes: Broca's, Wernicke's, anomic, conduction, transcortical motor, transcortical sensory and global aphasia syndromes. These syndromes are divided by three dichotomous criteria: fluency, comprehension, and repetition (Figure 1).

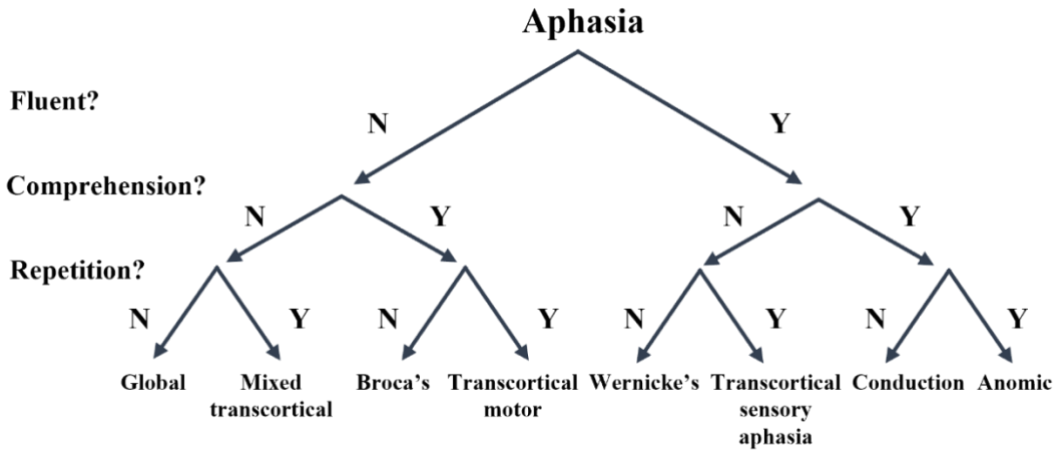


Figure 1. A flowchart of one the most influential aphasia classification systems. (Kaplan et al. 1983; Goodglass et al. 2001)

However, despite the depth and breadth of deficits in aphasia, almost all PWA experience some degree of word-finding difficulties (Dronkers and Baldo 2009) regardless of severity or location of lesion (Crinion & Leff, 2007). Word-finding difficulties ("anomia") describe when an individual knows the phonological, semantic, and orthographic properties of a word but has difficulty retrieving the word in the moment. These difficulties are often described as feeling a word is frustratingly irretrievable, but on the "tip of the tongue". Anomia can have a catastrophic adverse effect on daily life and decrease the likelihood of returning to work post-stroke (Graham, Pereira, & Teasell, 2011). Consequently, speech and language therapy (SLT) often focuses on word finding, increasing the ability to retrieve words in everyday life, and improving communication (Best et al. 2008).

SPEECH PRODUCTION

There is a consensus across the literature that the processing stages of speech production include: conceptual preparation, grammatical encoding, phonological encoding and monitoring of speech (Hillis 2002). These stages are exceedingly fast, in connected speech words are selected at a rate of two to three per second. Adult vocabularies are vast, storing tens of thousands of potential choice words. Although there is general

agreement regarding levels of processing, there is debate over the flow of information between these levels.

Discrete two-stage models

In two-stage models, the time course of lexical access includes an early semantic phase, followed by a later phonological phase. Amongst these models, there is a certain amount of shared architecture. Processing units are grouped into semantic, lexical, and phonological layers, whereby activation spreads through connected units.

Levelt (1989) proposed the influential discrete two-stage model of speech production. Various discrete processing components are sequentially activated in two-stage lexical access models. In the first stage, conceptual preparation involves generating the intent or message, considering the receiver. This leads sequentially to grammatical encoding, activating syntactic words ("lemmas") in the mental lexicon and creating an appropriate syntactic pattern and overall surface structure. The discrete and weight decay models use the term lemmas, whereas the semantic-phonological model terms these units 'word nodes'. However, the concepts are similar; the lemma level holds grammatical information.

The following proposed stage first involves morpho-phonological encoding. Here, it is suggested that the chosen form code is activated and the word's morphological and phonological configuration is accessed. This process is suggested to generate the 'phonological score', including the syllabified words, phrases and intonation pattern in its syntactic context. Phonetic encoding is then initiated, where each of the syllables in the phonological score trigger an articulatory gesture. Finally, the message is articulated through the laryngeal and supra-laryngeal apparatus, producing speech.

A key element of the discrete model is a complete separation between lemma selection and phonological encoding in a feedforward system. Connections between units are unidirectional, with no feedback of information from later to earlier levels. Further, information can only flow in serial order, with no cascading of information (McClelland 1979).

Interactive-activation two-stage models

In contrast, Dell (1986) proposed a spreading activation model, whereby the layers are interactive rather than discrete. In this model, the nodes of each adjacent layer are bi-directionally connected, with activation spreading interactively in both feedforward and feedback fashions. Dell and colleagues have since expanded upon the interactive model. Although Dell's (1986) model was initially produced to account for speech production in healthy individuals, the aphasia model (Dell et al., 1997) was proposed from implications of studies in PWA. This model was derived from similarity between healthy slips of the tongue and unsuccessful retrieval attempts in PWA, allowing for an explanation of error patterns in both groups.

The pattern of errors occurring in a PWA may inform the locus of impairment in the model (Halai et al. 2017). These errors can also be mapped within the brain using functional and structural neuroimaging. The aphasia model has five different categories of errors that occur during picture-naming tasks: semantic, formal, mixed, unrelated and non-word errors (Dell et al., 1997). Omission errors were initially excluded. However, these are very common in PWA and were later modelled (Dell, Lawler, Harris & Gordon, 2004). In response to the target item, a semantic error may occur, a meaning-related substitution by a word of the same category (cat → "dog"). Formal errors are real words that contain phonological segments (phonemes) of the target word (cat → "mat"). Mixed errors are real words that are related both semantically and phonologically (cat → "rat"). Unrelated errors are real words which are not semantically or phonologically related to the target (cat → "fog"). Non-word errors are responses that are not real words, with or without resembling the target (cat → "deg").

Individual differences in the number and type of errors made during picture-naming are simulated by differences in the associative strength of the bidirectional connections between layers. The broader hypothesis that patients' data can be captured by variations in the weight values in the model was used to develop two further model: the weight-decay (WD) model (Dell et al., 1997) and the semantic-phonological (SP) model (Foygel & Dell, 2000). Both these models feature cascading activation and feedback. Although these models are interactive with bidirectional connections, allowing for both feedforward and feedback activation, there is still a serial nature with separate lemma selection and phonological encoding stages.

The WD model assumes global impairments in either connection weights or activation decay (Dell et al., 1997). A global decrement in the model's connection weights, or increase in the activation decay rate, causes activation levels throughout the entire network to decrease. Lower activation levels can result in errors due to an increase in the signal to noise ratio. Changing the weight parameter leads to non-word and unrelated errors, whereas changing the decay leads to semantic, mixed, and formal errors.

The semantic-phonological (SP) model hypothesises that impairments stem from damage to lexical-semantic or lexical-phonological connections (Foygel & Dell, 2000). In this model, aphasia is assumed to alter the weight of connections in the lexical network, either between semantics/words or words/segments. Like the weight-decay model, there are two lesionable parameters. However, in this case, the global weight parameter is divided into a semantic weight and a phonological weight, and activation decay is no longer considered lesionable.

Schwartz, Dell, Martin, Gahl and Sobel (2006) performed a case series test of both models, finding both models to be successful in explaining patient variation on a task, with computational modelling of 94 PWA; however, the SP model explained the most variance. Abel, Willmes, and Huber (2007) compared the two model versions in model-orientated naming therapy. These results demonstrate the usefulness of both models regarding the effectiveness of potential therapies, but again, the SP model was better suited for model-orientated therapy. This paper concludes that the two-step interactive model complements the traditional box-and-arrow models used to guide therapy.

Distributed models

Parallel distributed processing, or PDP models (McClelland et al. 1986), have a different architecture to the previously discussed two-stage models. These models are constructed with neuron-like processing units, which mediate between input and output representations. These units transmit information of current activation status to other units in the model, through synapse-like connections. Units are grouped into layers and using gradual weight changes in connections between these layers, PDP models can be trained to map input to output using backpropagation (Rumelhart et al. 1986) or other learning algorithms.

One motivation for a PDP approach is the concept of graceful degradation. Although patterns of dissociations in neuropsychology are often striking, cognitive deficits are often graded. Damage does not generally result in a complete loss of a specific function, instead there is a process of graceful degradation. Damage causes graded, probabilistic deficits across multiple functions, with some sparing of performance. Thus, brain damage is not constrained to a single modular architecture. Therefore, PDP models are highly interactive, whereby activation levels of all units can influence any other unit in the model due to multiple interconnections between them. Damaging of connection weights within PDP models to simulate lesions results in graceful degradation of performance (Rumelhart et al. 1986).

The utility of these models is vast, with many applications including accounts of naming (Seidenberg and McClelland 1989), sentence comprehension (St. John and McClelland 1990) and lexical segmentation (cf. Davis, 2003), due to the ability of the models to learn new information and a neural inspired framework. However, the gradual process of weight changes means that mapping of novel representations can only occur in a slow process over many presentations. These novel presentations must be interleaved with existing mappings. If attempts are made to teach a model new information quickly, catastrophic interference occurs (McCloskey and Cohen 1989). This phenomenon refers to the tendency of an artificial neural network to completely forget previously learned information upon the learning of new knowledge.

Case-series and single case-study methodologies have also been used to explore different real-world applications speech production models. Lambon Ralph, Sage and Roberts (2000) explored two case studies with possible classical anomia (word-finding difficulties without semantic or phonological deficits) to a distributed model of speech production. A patient can theoretically "lose" a word in localist architecture if that lexical node is completely eradicated (Lambon Ralph, 1998). In this case, patients should never produce the word, even with the aid of priming or cueing. Whereas in a distributed architecture, a single item cannot be lost without affecting other elements due to the pattern representation of items over many elements. The PWA in this study could support this theory in that a proportion of lexical items appeared to be entirely lost. However, a graded hypothesis was proposed to account for consistently unnamed words and inconsistently named words. This proposal suggested that these items are not in fact lost, but rather have a very low probability of retrieval.

Neural theories of speech production

There are also various neural theories of speech production. Before the advent of functional neuroimaging techniques, researchers relied on patient studies of brain damage or electrical stimulation during neurosurgery to delineate neural structures necessary for function. Thus, traditionally models of language and speech production were ruled by the interplay of the posterior superior temporal lobe (Wernicke's area), the inferior frontal lobe (Broca's area) and the arcuate fasciculus, the white matter pathway connecting the two (Wernicke 1874). Lichtheim (1885) expanded upon the Broca-Wernicke models, adding the ventral language pathway. This white matter pathway included the extreme capsule and uncinate fasciculus. These tracts connect primary auditory and speech-motor areas (Lichtheim 1885).

With the proliferation of functional brain imaging studies, a consensus regarding the underlying processes involved in word production is emerging. The localisation of speech production still varies across participants; however, the use of lesion studies and neuroimaging has aided the identification of common areas of activity during specific tasks. Speech production processes are likely intrinsically interrelated with speech perception and language comprehension. There is extensive converging evidence of a dual-route model of speech processing (Hickok and Poeppel 2004, 2007). The bilateral ventral stream supports mapping from sound to meaning. Functional imaging studies implicate posterior middle temporal areas in this lexical-semantic processing (Rodd et al. 2005). The dorsal stream is proposed to map sound to action through auditory-motor integration (Hickok et al. 2000; Wise et al. 2001; Hickok and Poeppel 2004). A sensory-motor interface is necessary in speech production. The neural correlates of this mapping of sensorimotor integration implicate the left planum temporale region from neuropsychological (Anderson et al. 1999; Dronkers and Baldo 2009) and imaging (Hickok et al. 2000; Wise et al. 2001) evidence.

Indefrey and Levelt (2004) focused on commonalities across four tasks (picture naming, word generation, word reading and pseudoword reading) to identify reliably activated regions during distinct processing stages. During conceptual preparation, the visual stimulus of an image was underpinned by activation of the primary visual areas in the bilateral occipitotemporal regions and fusiform gyri. The inferior frontal gyrus (IFG) and the anterior/middle part of the middle temporal gyrus (MTG) were involved in

lemma selection. During phonological encoding, the left middle frontal gyrus (MFG), inferior temporal gyrus (ITG) and the superior temporal gyrus (STG) were activated. The ultimate stage of articulation was supported by activation of the pre-supplementary motor area, the left anterior insula and the left posterior pars opercularis.

There are practical issues to consider when imaging speech production. Talking during volume acquisition inevitably introduces motion-induced signal (Fiez 2001). To counter this issue, researchers often ask participants to speak like a ventriloquist, minimising head movement as much as possible. Covert speech has also been used to reduce this effect entirely. However, despite overlapping findings in overt and covert speech activation, covert speech is not entirely the same process as overt speech (Geva et al. 2011). An alternative is to use overt speech production during scanning, with rigorous preprocessing to remove motion artefacts.

Multi-echo fMRI allows for independent components denoising analysis, termed ME-ICA (Kundu et al. 2017). Dual echo paradigms have been previously shown to overcome temporal lobe signal loss by optimally combining echoes. It may seem illogical to use three echoes, as the third echo is of a length that is sub-optimal. However, three echoes allow for echo time (TE) relaxation curve estimation. Blood oxygenation level-dependent (BOLD) signal is dependent on the TE relaxation curve, whereas noise (or non-BOLD) is not. This estimation allows for the automatic or manual removal of non-BOLD signals and can mitigate motion artefacts beyond standard removal (Gonzalez-Castillo et al. 2016).

VOCABULARY ACQUISITION

Across the lifespan, languages evolve and change. New words emerge and old words take on new meanings. For example, a ‘tweet’ once described only the sound a bird made. A ‘tweet’ may now be a bird sound or a social media post. Word learning is an associative process, linking a phonological/orthographic form to a concept/external referent. There are multiple stages involved in word learning and different ways that word learning can occur.

For example, reading a novel word requires an abstraction from the visual form of a written word to access information about the spoken form of a word. A particular

functional region part of the left ventral occipitotemporal cortex (vOTC), an early visual processing region, has been labelled the "visual word form area" (Cohen et al. 2000). This region is associated with this lexical feedforward processing (Dehaene et al. 2002; Dehaene and Cohen 2011). There is evidence that the left vOTC is hierarchically graded along a posterior-anterior trajectory, along the visual ventral stream. Along this gradient, neural representations of visual letters become increasingly invariant to retinal location and transform to represent spoken language information (Taylor et al., 2019). In naturalistic language learning, novel words occur in the context of a sentence. This context may provide the necessary semantic information to garner meaning.

Rodriguez-Fornells et al., (2009) proposed an integrative model of spoken word learning in adults, containing three interfaces. The first is the dorsal-audio motor stream (Wise et al. 2001; Davis and Johnsrude 2007; Hickok and Poeppel 2007; Saur et al. 2008). This stream may support mapping of sounds to articulation, maintaining auditory-based representations of speech (Buchsbaum et al. 2005) and other sounds (Price et al. 2005). The second proposed stream is the ventral meaning integration interface, a proposal which aligns with the ventral language stream (Wise et al. 2001; Davis and Johnsrude 2007; Hickok and Poeppel 2007; Saur et al. 2008). The third stream is termed an episodic-lexical interface, based upon the theory that binding of new word representations should be medial temporal lobe (MTL) dependent (Squire et al. 2004; Davis and Gaskell 2009). This thesis focuses on this proposed episodic-lexical stream of associative word learning through word-referent pairings.

Complementary Learning Systems (CLS) model

In the late 1980s to early 1990s, several computational models were proposed with multi-layer artificial neural networks as discussed in General Introduction: Speech Production. Through backpropagation (Rumelhart et al. 1986), these models could acquire structured knowledge slowly with interleaved learning. However, if researchers attempted to teach these models new information quickly, catastrophic interference of previously acquired knowledge occurred (McCloskey and Cohen 1989). Humans can acquire new knowledge rapidly, with relative ease and without the forgetting of previously learned information. Thus, the CLS model (McClelland et al., 1995; Norman & O'Reilly, 2003) drew upon earlier ideas of Marr (1971) to provide a potential solution to the problem of catastrophic interference. The CLS model proposes that

multi-layer neural networks, similar to the neocortex of the brain need to be paired with a fast-learning, complementary system inspired by the human hippocampus.

This theory derives from various neuropsychological evidence centred on damage to the hippocampal system. Focal hippocampal lesions can produce a profound deficit in novel learning and memories, whilst other cognitive processes and memory performance is relatively spared. The prominent report of patient HM by Scoville and Milner (1957) first described this effect. Patient HM underwent bilateral hippocampal resection, with subsequent deficit of memory for events after or a few weeks prior to the lesion. In contrast, more remote memories remained intact.

A further important piece of neuropsychological evidence key to the predictions of the CLS model is temporally graded retrograde amnesia. This phenomenon was described by Ribot (1882). If an animal has an experience on the same day as bilateral hippocampal resection, that experience is near or totally lost from the individual's memory. However, suppose the hippocampi are removed after a period (one week, two weeks, or four weeks). In that case, the degree of recall increases indexed by the time between experience and onset of amnesia. This effect also occurs in human hippocampal amnesics who exhibit impaired recall of events prior to brain damage. The degree of deficit reduces as a function of time between event and injury (Ribot 1882).

The CLS theory proposes that the hippocampus and neocortex have specific complementary properties that allow for successful learning (McClelland et al. 1995). In this theory, the neocortex is viewed as a collection of partially overlapping processing systems. This system could be conceptualised as supporting performance of higher-level cognitive tasks through the transference of activation patterns between and within regions. The neocortex can thus perform slow learning of distributed representations, generalising across experiences. These representations are stable and can be retained for long periods. Conversely, the hippocampal system provides a mechanism for the rapid acquisition of pattern-separated experiences. Experiences or events may be represented by sparse patterns of activity, whereby individual neurons represent aggregations of the factors that lead to the event. Plasticity in the hippocampal-input connections increase the likelihood that further instances of elements involved in the prior event will elicit the full pattern activation and enable recall if the elements are sufficiently close to the stored pattern. This process may take multiple occurrences.

Reinstatement of pattern-separated experiences may also occur in sleep (Wilson and McNaughton 1994). Only a few studies have directly investigated hippocampal replay in humans during sleep (Axmacher et al. 2008; Schapiro et al. 2018) and waking unalert states (Carr et al. 2011). However, sleep-dependent knowledge consolidation is demonstrated in a vast amount of studies, with slow-wave sleep associated with improved recognition of new knowledge (Tamminen et al. 2010, 2013; Lewis and Durrant 2011; Gaskell et al. 2019).

CLS theory in novel word learning

Davis and Gaskell (2009) proposed that the general learning framework in CLS theory may apply to vocabulary learning. There are multiple sources of convergent evidence that suggest a role of the hippocampus in novel word learning. In an online fMRI associative learning study, Breitenstein et al. (2005) presented participants with an image and paired auditory pseudoword. Participants learned the novel vocabulary through associative learning exposure, with higher occurrences of "correct" pairings. There was strong evidence of initial hippocampal encoding of pseudowords. In addition, there was hippocampal modulation during online pseudoword learning, whereby a linear decrease of left hippocampal activity paralleled increases in pseudoword accuracy over the training.

Sleep also plays an essential role in the offline consolidation of new phonological word forms. A marker of word knowledge consolidation is competition between lexical representations (Gaskell and Dumay 2003). The process of learning to recognise a new word involves both the learning of the word itself, but also how to distinguish the representation from other words. When a novel pseudoword is newly learned ('cathedruke'), increasing levels of familiarity with the new word form is associated with slowed performance for existing, similar items (such as cathedral). However, lexical competition effects are only induced after a period of sleep. No lexical competition effects are observed if participants are awake for the same period (Dumay and Gaskell 2007).

Indeed, previous literature has demonstrated reductions in hippocampal-memory system involvement after a single night of sleep and consolidation of novel words. Davis et al. (2009) used fMRI to measure neural responses to novel pseudowords at different stages

of consolidation. Unfamiliar novel words had elevated hippocampal responses, and this response correlated with post-scanning measures of word learning. New word forms are initially stored separately from existing knowledge and slowly integrated over time. The novel word forms are first sparsely but rapidly encoded in the hippocampus. Offline consolidation during sleep strengthens distributed representations in neocortical memory systems (Davis and Gaskell 2009). The degree of performance change in word learning following sleep is associated with the frequency of slow-wave spindles (Tamminen et al. 2010) and the number of rapid eye movement (REM) periods (Thompson et al. 2021), supporting the proposal that sleep is a crucial step in memory consolidation of words.

There are multiple methods in the previous literature employed to investigate novel word learning. As adults have extensive vocabularies, it is a difficult task to find completely unknown, native and imageable items suitable across a group of participants. One method of novel word learning is to teach participants a second language (Raboyeau et al. 2004; Yang et al. 2015). Second language learning may be different to completely novel word learning. In non-native languages, there may be phonological elements that are unfamiliar to a non-speaker, either in inflections, trills or linkage of phonemes that increase the challenge of learning and recall. However, the learning of new names for existing items may cause issues. The target item has existing semantic and phonological representations in the lexicon, which may initiate proactive interference. Competition takes place between the two representations and skews accuracy and reaction time (Gaskell and Dumay 2003).

Researchers often use pseudowords to mimic native language. Pseudowords increase familiarity and decrease learning challenge for a more naturalistic language feel (Ozubko and Joordens 2011). Pseudowords are a powerful tool, using native phonemes to mimic natural language with control over psycholinguistic variables in completely novel word forms. However, pseudowords need to be paired with visual stimuli form to investigate word-referent learning. Some studies have used none real-world novel picture stimuli paired with pseudowords, such as “aliens from another planet” (Gupta et al. 2006). Arbitrary referents often do not have semantic meanings to aid in learning and consolidation. Takashima et al. (2017) trained participants with pseudowords, half with word meanings, to explore this issue. Participants completed a same-day and one-week later recognition fMRI task. Novel words with semantic information at encoding were

better retained. Regardless, many pseudoword studies demonstrate successful learning of novel items and provide second stage neocortical regions of interest, with differential responses to novel and existing words, including the left temporal lobe (Raboyeau et al. 2004; Davis et al. 2009), bilateral anterior temporal lobes (Grönholm et al. 2005) and fusiform gyrus (Breitenstein et al. 2005) with elevated responses during training.

In particular, Davis et al. (2009) presented participants with three sets of novel pseudowords without picture referents or associated semantic information. Two sets of items were intensively trained prior to scanning. Hippocampal activations were confined to entirely novel items during scanning, and these activations decreased with subsequent repetitions. The magnitude of this decline predicted performance on a test of recognition memory. Thus, these results provide clear evidence of the hippocampus involvement in the initial learning of novel words.

A neurotypical baseline of learning meaningful real-world items with native language names is needed to understand vocabulary re-learning in aphasia. A potentially suitable approach comes from a series of magnetoencephalography (MEG), positron emission tomography (PET) and aphasiological studies that used the 'Ancient Farming Equipment' (AFE) learning paradigm, which provides a line drawing, a novel Finnish name and a description of the item used (cf. Laine and Salmelin 2010). In the AFE protocol, participants learn native names for line drawings of real-world antique tools. The stimuli tools have become obsolete and thus are unfamiliar for most participants.

Cornelissen et al. (2004) conducted the first experiment with the AFE paradigm, measuring cortical effects of novel word learning using magnetoencephalography (MEG). Five healthy participants underwent MEG before, during and after learning the novel items. Half of the items were trained with phonological information (item name) and half with semantic information (tool use). A learning effect was observed in the inferior parietal lobe (IPL) for three out of the five participants.

Pohl et al. (2017) used pseudowords to explore functional brain organisation in healthy participants before, during and after standardised speech production training using fMRI and a learning paradigm derived from behavioural treatments for aphasia patients (Abel, Weiller, Huber, Willmes & Specht, 2015). Participants were trained to learn pseudoword names for existing items. In contrasts of early training-related activation and consolidated learning activation, there was increased activation in areas involved in

monitoring (anterior cingulate cortex) and selection between multiple lexical components (bilateral caudate), which aligns with the response selection theory. In PWA, therapy-induced recovery was associated with abnormal monitoring functioning and decreased bilateral caudate activation.

A note to consider is that any review of the fMRI literature tends to omit ventral anterior temporal lobe (vATL) regions due to susceptibility-induced signal loss in the temporal lobes (Devlin et al. 2000). fMRI is susceptible to distortion artefacts caused by inhomogeneity in the magnetic field where tissues changes occur, such as from brain tissue to sinus cavity, with the altered field causing signal breakup and misplacement. Visser, Jefferies, & Lambon Ralph (2010) conducted a meta-analysis of functional neuroimaging studies with semantic tasks. Four factors influenced the likelihood of finding ATL activation in semantic processing studies: using PET rather than fMRI, a field of view of more than 15cm, using a high-level baseline task and the inclusion of the ATL as a region of interest. In terms of aphasia research, enhanced protocols can be used to ensure the inclusion of ATL activation if there is indeed any, as suggested by the results of bilateral ATL activation in semantic processing. An example of these enhanced protocols is a dual-echo fMRI approach, which has been shown to have a higher sensitivity in the ATL (Poser et al. 2006; Halai et al. 2014, 2015a). Therefore, it is important to select an enhanced protocol and *a priori* ROIs to have the required level of sensitivity to observe BOLD activity in this region.

VOCABULARY RE-ACQUISITION

Speech and language therapy

A pervasive feature of post-stroke aphasia is anomia. Speech and language therapy (SLT) thus often focuses on word-finding, increasing the ability to retrieve words in everyday life, and improving communication (Best et al. 2008). It is unknown whether PWA re-acquire or re-learn vocabulary post-stroke, with debate as to whether aphasia is a disorder of storage or access (Shallice 1988). In a storage disorder, items cannot be consistently retrieved, and priming is ineffective. This means that the mental representations are lost. In a disorder of access, items are inconsistently retrieved, and

priming is effective. In this case, the mental representations are temporarily not accessible (Shallice 1988). Phonological and semantic priming techniques are often effective in SLT (Abel et al., 2003; Conroy et al., 2009; Fillingham et al., 2006; Nickels, 2002). However, there is vast variability in treatment outcome and performance within post-stroke aphasics. Factors beyond variable lesion profile and language deficits may influence treatment outcomes. For example, better recognition memory, executive functioning and monitoring ability have been associated with improved response to cueing therapy (Fillingham et al. 2006).

Speech therapy can be ineffective for some patients or provide less successful results. There are no explicit measures of which patients would benefit from which treatment. This highlights the need to understand better the neurobiological mechanisms associated with aphasia intervention. Researchers have, however, identified vital regions of the left hemisphere such as Brodmann's areas 37 and 39 which, when lesioned, are associated with fewer improvements in naming (Fridriksson 2010). Furthermore, global and left temporal structural connectivity preservation accounts for variability in treatment-related naming improvements in aphasia (Bonilha et al. 2016). Further research in this area may provide a measure for practitioners to better predict patient language outcomes post-treatment.

The neural processes driving SLT outcomes are also unclear. SLT may induce standard mechanisms of recovery, of which there are three overarching theories: 'perilesional', 'laterality-shift' and 'disinhibition' hypotheses. The 'perilesional' hypothesis proposes that recovery is dependent on the reconstitution of language systems in undamaged tissue around the lesion (Hillis et al. 2006; Saur et al. 2006; Meinzer and Breitenstein 2008; Fridriksson 2010; van Oers et al. 2010). Abel et al., (2015) demonstrated, in a group of 14 patients with post-stroke aphasia, positive SLT outcomes were associated with spared left inferior frontal gyrus (IFG) activation during naming across both phonological and semantic therapies. These results were also found by Van Hees et al. (2014) in a phonological treatment study of eight patients, whereby treatment-induced improved naming was associated with greater perilesional activity.

The 'laterality-shift' hypothesis proposes that aphasia recovery can be attributed to a 'shift' of language function to the nondominant contralateral (typically right) hemisphere (Blasi et al. 2002; Leff et al. 2002; Winhuisen et al. 2005; Turkeltaub et al.

2012). Tuomiranta et al. (2014) reported a patient with extensive left-hemispheric damage that responded to orthographic naming therapy. Improved behavioural performance was associated with right hemispheric temporal and parietal regions, supporting the 'laterality-shift' hypothesis. The 'disinhibition' hypothesis is the proposal that when the dominant left hemispheric language system is damaged the errorful right hemispheric system is released from inhibition (Naeser et al., 2004, 2005; Rosen et al., 2000; Thiel et al., 2006). However, best recovery may be associated with the dominant and more effective left hemispheric systems taking over once more and subsequently reinhibiting the less effective right hemispheric systems (Heiss and Thiel, 2006; Naeser et al., 2005; Rosen et al., 2000).

The non-converging theories of the neural correlates supporting aphasia recovery may be due to the heterogenous nature of post-stroke aphasia. Abel et al. (2015) investigated intervention-induced changes in brain activation, using increasing semantic and phonological cueing-hierarchies. There was a differentiation in brain activation between phonologically-impaired (P) patients and semantically-impaired (S) patients. P-patients relied on preserved functional brain areas in the left hemisphere. In contrast, S-patients relied on right hemispheric compensation in the right inferior frontal gyrus, pars triangularis, as well as bi-hemispheric strategies. In group level studies, these varying profiles of deficits and lesions are averaged, which may eradicate effects.

SLT may also drive neurotypical processes. These processes may be those intrinsic to speech, such as verbal short-term memory (Peñaloza et al. 2016; Coran et al. 2020), or processes more generally related to knowledge acquisition. Conroy, Sage and Lambon Ralph (2009) found that visual working memory (WM) was a significant predictor of therapy outcome in picture naming. STM and WM differ in that although both are limited capacity cognitive storage systems, WM allows for the manipulation of stored information, whereas STM does not. It appears that memory and other non-linguistic cognitive factors play a role in vocabulary acquisition post-stroke, and these processes may need to be assessed and trained during therapy.

Overall, patients respond differentially to therapy, and there is currently a poor understanding of what predicts these individual differences. Additionally, there is conflicting evidence within the literature regarding whether PWA employ the same neural mechanisms post-therapy as healthy participants when word-finding. For

example, Léger et al. (2002) found that the patterns of activation in PWA were similar to controls after word-finding therapy. In contrast, Fridriksson et al. (2008) found that patterns of activation in PWA were not similar to controls. This disparity highlights that the neural underpinnings of therapy and intervention in post-stroke aphasia remain a gap in knowledge. Understanding the underlying neural mechanisms involved in treatment may allow tailored treatments and better outcome predictions.

Spontaneous recovery

Most people with post-stroke aphasia experience spontaneous entire, or partial, recovery in the initial recovery phase, even without intervention. Many research studies have evidenced this. For example, Hartman (1981) tested 44 patients 14 days post-onset, then 30 days later. 41 of the 44 patients experienced significant improvement without any therapy. Lendrem and Lincoln (1985) demonstrated spontaneous recovery between 4 and 34 weeks in aphasic patients.

Neuroplasticity may account for spontaneous recovery. Welbourne and Lambon Ralph (2005) used a PDP model of reading to model the effects of spontaneous recovery using a plasticity-based re-learning process. They found that recovery was faster than initial training and that full recovery was possible with neuroplasticity alone. The first two to three months following stroke are crucial for spontaneous neuroplasticity. There is natural neurophysiological repair and cortical reorganisation of function during this period (Robertson & Fitzpatrick, 2008) and often rapid progression of language ability (Lazar et al. 2008). Ueno, Saito, Roger and Lambon Ralph (2011) developed a neurocomputational dual dorsal-ventral model of language pathways. Simulations of this model suggested that the division of labour between the two pathways is plastic. If one pathway is damaged, processing can be re-optimised within and between the pathways. This plasticity may mimic the spontaneous recovery often observed post-stroke (Lambon Ralph, Snell, Fillingham, Conroy, & Sage, 2010; Welbourne & Lambon Ralph, 2007).

Spontaneous recovery is an important consideration when assessing therapy efficacy, as some recovery may be spontaneous and occur without intervention. However, this only occurs in the acute phase, with decreasing evidence of spontaneous recovery in the post-

acute phase. Therefore, a key inclusion criterion in this thesis was chronic rather than acute aphasia.

Novel word learning

To fully understand whether patients are re-acquiring or re-learning vocabulary, it may be necessary also to study the differences and similarities between healthy adult and patient word acquisition. To date, few PWA studies have used novel words for therapy, and fewer have explored the neural correlates of novel word learning in aphasia. This may be due to the assumption that individuals would be unable to learn novel words. Previous research has shown that phonologically-impaired PWA do show poor performance on non-word tasks across many domains, including reading, repetition and phonological manipulation (Jefferies et al. 2006). Therefore, a pseudoword target in post-stroke aphasia may prove difficult. Alternatively, PWA may prefer to practice words that can be used in everyday life, as there may be no or limited generalisation from therapy to untreated items (Abel, Schultz, Radermacher, Willmes, & Huber, 2005; Best et al., 2013; Conroy, Sage, & Ralph, 2009; Fillingham, Sage, & Lambon Ralph, 2006). However, novel native words provide a more stimulating and challenging task for individuals with relatively mild naming deficits.

The AFE paradigm, as discussed in General Introduction: Vocabulary acquisition, has been used to explore novel word learning in a small number of PWA. Tuomiranta, Rautakoski, Rinne, Martin, and Laine (2012) explored the long-term maintenance of words learned through the AFE paradigms in two participants with chronic non-fluent aphasia, finding that one individual retained 50% of the items six months post-training, whereas the other retained 25%. These results demonstrate that PWA can name novel items and maintain some of this learning. The ability of PWA to learn and retain the novel AFE items is associated not only with the severity of disorder, but also with verbal STM capacity (Peñaloza et al. 2016). Tuomiranta et al., (2015) studied a patient with an extensive left hemispheric lesion and disconnection of the arcuate fasciculus. Despite high levels of damage and moderate aphasia, this patient learned and maintained novel AFE items on par with healthy controls up to 6 months post-learning. The patient and healthy control participants showed increased hippocampal activation while learning novel versus familiar word-picture pairs, indicating that neurotypical

word learning and word re-learning in post stroke aphasia may have similar neural correlates.

KEY THESIS AIMS

Multiple theoretical and clinical issues warrant further exploration based on the previously discussed literature. By taking a global view of the current literature, these issues can be coalesced into four key research aims that will be the focus of this thesis. There is an over-arching thread of combining SLT-style word training with neuroimaging to improve our understanding of vocabulary acquisition, consolidation and retention in neurotypical and patient populations. The central research aims were as follows:

1. Investigate whether word learning in healthy populations fits within both stages of the CLS model, with initial learning supported by the hippocampal-episodic memory systems, followed by a gradual shift in the division of labour to domain-specific neocortical language areas.

2. Explore the levels of interactivity between episodic and language-semantic systems at differing stages of consolidation. What are the timescales involved in this shifting division of labour? Is there a permanent role of episodic activity in picture naming, or is this activation time-limited to consolidatory processes?

3. Examine whether the set of ideas termed variable neuro-displacement are applicable beyond recovery in post-stroke aphasia to neurotypical language learning. Does bilaterality of word learning systems in adults provide resilience to increased performance demands?

4. Does re-learning or novel word learning in post-stroke aphasia fit within this healthy learning framework? If so, is there a level of critical damage to the dominant left hemispheric language regions, resulting in deviations from the norm? If not, what are the processes that drive treatment outcomes in picture naming treatment?

Acknowledgement of authors' contributions

Professor Matt Lambon Ralph and Dr Anna Woollams supervised all Chapters of this thesis. Dr Steffie Bruehl also supervised Chapter 1 of this thesis. I recruited all patients from the Neuroscience and Aphasia Research Unit (NARU) database, where background assessments and structural scans are available, and all healthy participants from the community in the North West. I also recruited additional patients to the NARU database from stroke support groups, thus undertaking the preliminary neuropsychological testing and structural scanning. With guidance from my supervisors, I collated the stimuli and made the word learning and word re-learning program. Dr Ajay Halai provided guidance on the triple echo fMRI parameters and preprocessing for Chapter 1. I analysed the effectiveness of a ME-ICA paradigm versus optimally combined triple echo and dual-echo paradigms. I collected, preprocessed and analysed all behavioural and neuroimaging data in this thesis. I wrote the initial draft for each Chapter of this thesis, which were further developed with feedback from my supervisors.

CHAPTER 1: DIRECT NEURAL EVIDENCE FOR THE CONTRASTIVE ROLES OF THE COMPLEMENTARY LEARNING SYSTEMS IN ADULT ACQUISITION OF NATIVE VOCABULARY

ABSTRACT

The Complementary Learning Systems (CLS) theory provides a powerful framework for considering the acquisition, consolidation and generalisation of new knowledge. We tested this proposed neural division of labour in adults through an investigation of the consolidation and long-term retention of newly-learned native vocabulary with post-learning functional neuroimaging. Newly-learned items were compared to two conditions: (i) previously known items to highlight the similarities and differences with established vocabulary; and (ii) unknown/untrained items to provide a control for non-specific perceptual and motor-speech output. Consistent with the CLS, retrieval of newly-learned items was supported by a combination of regions associated with episodic memory (including left hippocampus) and the language-semantic areas that support established vocabulary (left inferior frontal gyrus and left anterior temporal lobe). Furthermore, there was a shifting division of labour across these two networks in line with the items' consolidation status; faster naming was associated with more activation of language-semantic areas and lesser activation of episodic memory regions. Hippocampal activity during naming predicted more than half the variation in naming retention six months later.

Published paper: Gore, K. R., Woollams, A. M., Bruehl, S., Halai, A. D & Lambon Ralph, M. A. (in press). Direct neural evidence for the contrastive roles of the complementary learning systems in adult acquisition of native vocabulary. *Cerebral Cortex*.

INTRODUCTION

Across the lifespan, humans need to acquire new knowledge and do so rapidly with relative ease. One lifelong learning process is vocabulary acquisition. Beyond the initial influx of new language in childhood, there are numerous words, meanings and expressions to learn throughout adulthood. Thus, individuals constantly acquire new vocabulary relating to their everyday lives, hobbies and profession. Re-establishing vocabulary is also a key target for those with language impairment (aphasia) after brain damage from injury, stroke or dementia because word-finding difficulties (anomia) are a pervasive and frustrating feature of all types of aphasia (Rohrer et al. 2008). Therefore, from both cognitive and clinical neuroscience perspectives, it is fundamentally important to understand both the cognitive and neural bases of vocabulary acquisition.

One influential theory is the Complementary Learning Systems (CLS; Marr, 1971; McClelland, McNaughton, & O'Reilly, 1995) model. This theory proposes that new knowledge is initially coded through rapidly formed, sparse representations supported by the medial temporal lobes (MTL) and hippocampus. Longer-term consolidation and evolution of generalisable representations follow from slower, interleaved learning and MTL replay to neocortical regions. Thus, over time, there is a gradual shift in the division of representational load between MTL and neocortical regions (with the rate of change depending on various factors: cf. McClelland et al., 2020). The CLS provides a potentially generalisable theoretical framework for the acquisition of many different kinds of knowledge including language acquisition (cf. Davis and Gaskell, 2009). There is, however, little direct neural evidence for this theory in long-term language learning, particularly in adults who already have large and varied vocabularies.

To date, few if any studies have explored the processes that underpin new vocabulary learning within adults' native language (i.e., learning the meaning and name of novel items/concepts as one might do when learning about a new hobby, profession, or technology). Instead, the handful of pre-existing investigations have typically focused on second language learning. Studies have adopted different experimental designs. Some have required participants to link brand new names to pre-existing, well-established meanings (Raboyeau et al. 2004; Yang et al. 2015). Alternatively, to avoid the unfamiliar phonetic and phonological elements of second languages, researchers have used pseudowords that conform to the phonological structure of the native

language (Mestres-Missé et al. 2008; Davis et al. 2009; Paulesu et al. 2009; Ozubko and Joordens 2011; Pohl et al. 2017). Pseudowords, however, do not have semantic meanings to aid learning and consolidation. Takashima et al. (2017) trained participants with pseudowords, half with word meanings, to explore this issue. Participants completed a same-day and one-week later recognition fMRI task. Novel words with semantic information at encoding were better retained, but utilised both the episodic and semantic systems during recognition at both stages. Of course, learning additional names for pre-existing items may generate competition between new and old words when naming. This proactive interference can skew accuracy and reaction times (Gaskell and Dumay 2003). To avoid these issues, researchers sometimes use abstract (i.e., meaningless) images alongside pseudowords (Takashima et al. 2014).

Although these pseudoword studies assessed performance through recognition tasks rather than the full recall process needed in speech production, they nevertheless indicate some important target brain regions for investigating native vocabulary. In an online fMRI associative learning study, Breitenstein et al., (2005) presented participants with an image and paired auditory pseudoword. Participants learned the novel vocabulary through associative learning exposure, with higher occurrences of "correct" pairings. There was strong evidence of initial hippocampal encoding of pseudowords. In addition, there was hippocampal modulation during online pseudoword learning, whereby a linear decrease of left hippocampal activity paralleled increases in pseudoword accuracy over the training. Davis et al. (2009) used fMRI to measure neural responses to novel pseudowords at different stages of consolidation. Unfamiliar novel words had elevated hippocampal responses and this response correlated with post-scanning measures of word learning. Similar to Breitenstein et al. (2005), as participants completed more training, there were associated hippocampal activity decreases. These studies provide evidence for the first stage of the CLS, in short-term learning of pseudowords. Such studies also provide second stage neocortical regions of interest, with differential responses to novel and existing words, including the left temporal lobe (Raboyeau et al. 2004; Davis et al. 2009), bilateral anterior temporal lobes (Grönholm et al. 2005) and fusiform gyrus (Breitenstein et al. 2005) with elevated responses during training.

To fully understand native vocabulary acquisition and recovery of vocabulary in aphasia, investigation of meaningful real-world items with native language names

would be ideal. A potentially suitable approach comes from a series of MEG and aphasiological studies that used the 'Ancient Farming Equipment' learning paradigm, which provides a line drawing, a novel Finnish name and a description of how the item is used (cf. Laine and Salmelin 2010). However, to fully elucidate the networks supporting word acquisition, and allow charting of the neocortical transfer proposed by the CLS, a longer-term strategy is required.

In the present study, we generated a direct evaluation of the CLS with respect to native vocabulary acquisition, including the role of semantic learning. Accordingly, we used fMRI to investigate the interaction between episodic and semantic neural networks that underlie native novel vocabulary learning, and how these processes differ to long-standing fully consolidated words. Healthy, older participants were recruited for comparability with aphasic patient samples, and due to increases in word-finding difficulties in normal ageing (Burke & Shafto, 2008). Participants were trained on novel native words for three weeks, before performing both picture naming (of previously known items, untrained/unknown items and select trained items which had been learned successfully per participant) and semantic judgement tasks in the scanner (i.e., names had to be learned sufficiently well for speech production rather than simply above-chance memory recognition). We also adopted this method and learning target as it directly mimics those found in rehabilitation of aphasic word-finding difficulties (where patients aim to re-establish meaningful, native vocabulary through multiple learning sessions, extending over several weeks). Consequently, not only does the current study provide information about native vocabulary acquisition in the healthy brain, but it may also give important clues about the neural bases of successful aphasia rehabilitation by providing a baseline for the same analysis in patients with aphasia.

We predicted that at a whole brain level, naming of newly trained, less consolidated words (for a maximum of three weeks e.g., echidna, dilruba, binnacle) would rely on the episodic/MTL areas as described by the first stage of the CLS. Whereas naming of already known, highly consolidated words (e.g., dragonfly, xylophone, hairdryer) would rely on the language network, i.e., the neocortical second stage of the CLS. We used behavioural measures of naming accuracy and reaction times (RT) to measure how well learned and consolidated the newly learned items were. For the newly-learned vocabulary, we predicted that there would be a positive correlation between regions of interest in the episodic network, namely the bilateral hippocampi and left inferior

parietal lobe, and longer RTs (i.e., for items that were not as well consolidated). We predicted the opposite would occur with regions of interest in the language network, with more BOLD activity in these regions correlating with quicker RTs (i.e., reflecting the gradual shift from episodic/MTL regions to neocortical ones for the most consolidated items). In contrast, for naming the established items, BOLD activity within the MTL/episodic regions would not be expected to have any significant correlations with performance measures, as this vocabulary should be well consolidated into the language system and thus rely on the language network alone. Finally, we considered the relationship between initial consolidation efficacy with longer-term retention of the newly-acquired vocabulary. Specifically, we tested the hypothesis that the items which were less well consolidated after initial learning (as indexed by their higher reliance on the MTL/episodic network) would be less well retained after six months, whereas items that were better consolidated (as indexed by their higher activation of the language network) would be better retained.

In this study, we explored the following questions: i) does vocabulary acquisition follow the CLS framework of learning? ii) does involvement of the episodic system when naming newly trained words correlate with worse performance, and does involvement of the semantic-language system when naming newly trained words correlate with better performance? iii) If so, do these correlations significantly differ from naming previously known items?

METHODS

Participants

Twenty older, healthy native-English speakers were recruited (twelve females, age range 46-76 years, mean age 63.90, *SD* 8.82). All participants were right-handed, with normal or corrected-to-normal vision, no history of neurological disease, dyslexia or contraindications to MRI scanning. The Addenbrooke's Cognitive Examination Revised (ACE-R) was used to screen for dementia, with a cut-off score of 88. Capacity for verbal learning was tested with the California Verbal Learning Test (CVLT). All participants gave informed consent before participating and the study was approved by a local National Health Service (NHS) ethics committee.

Stimuli

There were three sets of stimuli items (for a full list, see Supplementary Table 1). All sets contained real-world items including mammals, fish, birds, tools, food, clothing and toys. Two sets included unfamiliar items with very low word frequency names. These items were drawn from the British National Corpus (BNC; Davies, 2004), a 100-million-word text corpus. One set was used for training, whilst the other remained as an untrained baseline set. The trained and untrained sets were counterbalanced across participants. The third set contained familiar items. These items were drawn from the International Picture Naming Project. Items were selected that could be named accurately (85-100%), with low word frequency and longer reaction times (>1000ms) to select less easily named items. All stimuli were below a word frequency of 100 words per 100-million and had high name agreement. For the baseline task, the item images for the known, trained and untrained sets were phase scrambled. In the picture naming task, fMRI stimuli were single high quality, coloured photographs with a white background. In the semantic decision task, the fMRI stimuli were presented as an orthographic written name, in black text on a white background.

Procedure

There were five stages: baseline naming assessment, word training, post-training behavioural assessment, functional imaging data collection and maintenance naming assessment (Figure 1). Participants were tested on all items before training. Stimuli sets were tailored to each participant so that all known items could be named, and all untrained and to-be-trained items could not be named prior to training. Participants undertook fMRI scanning within 2 days of finishing training. Only items which had been successfully learned, demonstrated in the post-training naming assessment, were used in the fMRI trained condition (therefore there were different stimuli sets per participant for the trained condition). To assess maintenance, participants were tested on learned items between five- and six-months post scanning, without interim training.

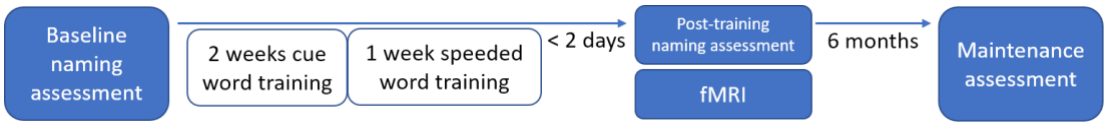


Figure 1. Timeline of study stages.

Behavioural training

Participants received self-guided, at-home training on new words and the related semantic information. Training took place for up to 45 minutes a day, four days a week for three weeks. In the first two weeks, participants received cue training. In the third week, participants received speeded training.

Items were presented via an interactive PowerPoint presentation. Visual Basic for Applications was used to store cue choice, time on task and accuracy data. In weeks 1 and 2, cue training took place daily. A novel picture was shown, with the name both in orthographic and audio form. Participants were instructed to listen to the name and repeat it out loud. After all items had been repeated the cue training began. Participants were instructed only to use cues when they needed one and reminded they would be tested on the semantic information. The training was designed to allow healthy participants to choose the level of cue they thought they would need to be correct on each trial. This interactive and self-determined approach allowed the training to feel challenging, engaging and reduce boredom.

The cue training was commonly used in standard speech and language therapy (Nickels 2002; Abel et al. 2005; Pohl et al. 2017). Participants saw a picture of an item with a choice of four cues, or the option to name the item with no cues. Participants could use as many cues as they like, in any order. There were four increasing cues. First, a picture plus a written descriptive semantic cue. Second, the picture plus the first name phoneme. Third, the first and second name phoneme were cued. The fourth cue was the whole name. All cues were given both orthographically and audibly. The semantic cue was formed in the same way for each item, initially with the geographical origins, then an identifying feature, followed by a broader semantic cue. For example, an ankus was "An Indian hooked tool used to handle and train elephants."

After each naming attempt, the whole correct word was given. Participants were asked to indicate whether they named each item correctly or not. Participants then indicated whether the item was European or not. The initial training set was 10 items. When participants were able to name 70% of the presented items with no cue, then another 10 items were added to the set, incrementally up to 50 items.

In the third week of training, the learned items were used in a novel repeated increasingly-speeded presentation (RISP; Conroy, Sotiropoulou Drosopoulou, Humphreys, Halai, & Lambon Ralph, 2018) learning environment. Participants were instructed that the computer would present an item for a short time, and they needed to name the picture before a specified time limit. When participants reached a success rate of 70% at a target speed, the timing was incrementally reduced from 1.8s to 1.4s, to 1s. When participants beat the 1s target for 70% of items, the set size was increased by 10 items and the timing was reset to 1.8s.

We assessed participants' learning using a post-training assessment of trained items in the absence of cues. Only successfully named items were used during the fMRI naming task (trained vocabulary condition; M = 45 items), creating participant-specific trained condition naming sets. The fMRI session took place on the same day as the post-training assessment.

Neuroimaging acquisition

All scans were acquired on a 3T Phillips Achieva scanner, with a 32-channel head coil with a SENSE factor of 2.5. High resolution, whole brain, structural images were acquired including 260 slices with the following parameters: TR = 8.4ms, TE = 3.9ms, flip angle = 8 degrees, FOV = 240 x 191mm, resolution matrix = 256 x 206, voxel size = 0.9 x 1.7 x 0.9mm.

We opted to use a triple gradient echo EPI sequence in order to improve the signal-to-noise ratio (SNR), particularly in the anterior temporal lobes (ATLs) where traditionally there are issues of EPI signal dropout and distortion (Poser et al. 2006; Halai et al. 2014, 2015). All functional scans were acquired using an upward tilt up to 45 degrees from the AC-PC line to reduce ghosting artefacts from the eyes into the temporal lobes. The sequence included 31 slices covering the whole brain with TR = 2.5s, TE = 12, 30 and 48ms, flip angle = 85 degrees, FOV = 240 x 240mm, resolution matrix = 80 x 80, and voxel size = 3.75 x 3.75 x 4mm.

All stimuli were presented electronically using E-Prime 2.0 software (Psychology Software Tools). The block order was pseudo-randomised optimised for statistical power using OptSeq (http://surfer.nmr.mgh.harvard.edu/optseq/). Verbal responses were recorded using a fibre optic microphone for fMRI (FOMRI; Optoacoustics) with noise-

cancelling. Participants were instructed to speak ‘like a ventriloquist’ to reduce motion artefacts.

Participants performed two tasks during imaging acquisition, one of which is the focus of a separate study. For this study, a picture naming task comprised a block design with four conditions; known, trained, untrained and baseline. In the known condition, participants overtly named familiar items (e.g., umbrella). In the trained condition, participants named newly learned items (e.g, echidna). If participants could not remember an item name they responded: “don’t know”. If the item was novel (untrained condition), participants also responded “don’t know”. Similarly, participants responded “don’t know” to phase-scrambled stimuli from the other conditions as the baseline condition. The task included two trial speeds but the results did not differ across these conditions, therefore data were collapsed across this manipulation. In the standard speed condition, each 1900ms trial consisted of a fixation cross for 700ms, followed immediately by the target image in the middle of a white screen for 1200ms. With 5 items per block, each block lasted 9.5s. We also included 8 rest blocks per run which were jittered to have an average length of 9.5s. With 32 task blocks and 8 rest blocks per run, the total run time was 6 minutes and 33 seconds. In the slower condition, each trial lasted 3700ms and consisted of a fixation cross for 700ms, followed by the target image for 3000ms. Only 3 items were presented per block and each block lasted 11.1s. As before, 8 jittered rest blocks were included with an average length of 11.1s. With 32 task blocks and 8 rest blocks, the total run time was 7 minutes and 4 seconds.

The second task, the focus of a separate study, required participants to make semantic decisions. This included three blocked conditions; trained, untrained and baseline. In the trained and untrained condition, participants responded “Yes” or “No” or “Don’t Know” to the semantic question “Is it European?”. In the baseline task, participants responded “Up” to an ascending alphabetical sequence “ABCD” or “Down” to a descending alphabetical sequence “DCBA”. As above, we used two trial speeds but found no differences between conditions; therefore, data were combined. In the standard speed condition, a fixation cross was displayed for 700ms, followed immediately by the target image for 1200ms (total trial = 1900ms). There were 5 trials per block each lasting 9.5s, with 6 jittered rest blocks averaging to 9.5s. The total run time was 6 minutes and 33 seconds, which included 24 task and 6 rest blocks. In the slower condition, displayed the target image for 3000ms (total trial = 3700ms). A total of 24 task blocks were used

with 3 trials per block (11.1s) and 6 jittered rest blocks averaging to 11.1s (total run time = 7 minutes and 4 seconds).

Neuroimaging pre-processing and analysis

T1 data was pre-processed using the FMRIB Software Library (FSL, version 6.0.0; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/, Woolrich et al., 2009). Brain tissue was extracted from the structural images (BET; Smith, 2002) and an initial bias-field correction was applied using FSL's anatomy pipeline (FAST; Zhang, Brady, & Smith, 2001), excluding sub-cortical segmentation as this was performed with BET. Registration to standard space was performed in FSL with FLIRT and FNIRT (Woolrich et al. 2009; Patenaude et al. 2011) and segmentation with FAST (Zhang et al. 2001). Despiking and slice time correction was applied to the functional data in the AFNI neuroimaging suite (v19.2.10; Cox, 1996; Cox & Hyde, 1997; 3dDespike; 3dTshift). Combined normalisation, co-registration and motion correction parameter sets were applied to each functional echo in FSL. Functional data were optimally combined, taking a weighted summation of the three echoes, using an exponential T2* weighting approach (Posse et al. 1999) and regression analysis. Functional runs were also combined and denoised using multi-echo independent component analysis (ME-ICA; Kundu et al., 2013; Kundu, Inati, Evans, Luh, & Bandettini, 2012) using the tool meica.py (v3.2) in AFNI (Cox, 1996; Cox & Hyde, 1997). The denoised timeseries were normalised to standard space using FNIRT warps, then smoothed.

Statistical whole brain and region of interest analyses were performed using SPM12 (Wellcome Trust Centre for Neuroimaging) and MarsBaR. Regions of interest (ROIs) were based upon previous literature. Medial temporal lobe structures, including bilateral hippocampi are critical for episodic memory, as evidenced by hippocampal amnesia (Dickerson & Eichenbaum, 2010). However, episodic memory processes also involve the inferior parietal lobe (IPL), despite parietal lesions not resulting in episodic memory deficits (Cabeza et al. 2008). The left inferior frontal gyrus (IFG) is considered critical in speech production and semantic processes (Blank et al. 2002a; Hickok and Poeppel 2007; Lazar and Mohr 2011; Price 2012). The middle temporal gyrus (MTG) is activated during semantic processing (Binder et al. 2009; Visser et al. 2012; Noonan et al. 2013; Jackson 2021), and focal damage is associated with semantic deficits (Dronkers et al. 2004). The specific co-ordinates for these ROIs were derived by

conducting a Neurosynth (Yarkoni et al. 2011) fMRI meta-analysis using two search terms: 'episodic memory' (bilateral hippocampi; MNI: -28 -14 -15, 29 -14 -15 and left IPL MNI: -47 -64 34), and 'language' (left IFG MNI: -46 28 10 and left MTG MNI: -52 -42 0).

Furthermore, we included a left ventral anterior temporal lobe (vATL) ROI (MNI: -36 -15 -30) taken from a key reference (Binney, Embleton, Jefferies, Parker, & Lambon Ralph, 2010). The vATL is often missed in fMRI studies using typical echo times of >30ms at 3T due to signal dropout. However, there is clear evidence from the neuropsychology literature and semantic dementia patients that the vATLs are important for semantic cognition patients (Rogers et al. 2004; Patterson et al. 2007; Lambon Ralph 2014; Lambon Ralph et al. 2016). Indeed, there is growing evidence that fMRI protocols optimised for signal detection in areas of magnetic susceptibility can identify vATL areas during semantic processing (Devlin et al. 2000; Halai et al. 2014, 2015; Jackson et al. 2016; Rice et al. 2018).

Mean beta values were extracted from each ROI. Reaction times (RTs) for ROI analyses were calculated from onset of stimulus and were z-scored to account for any variance due to time on task. RTs were z-scored by condition to enable analysis of within-condition RT variance.

RESULTS

Behavioural Data

Participants spent a mean of 4.3 hours training (SD = 0.8) over an average of 12 sessions. Participants successfully learned the novel vocabulary with an average gain of 81% (SD = 10.73) outside the scanner. Inside the scanner, in the trained condition, participants were presented with only items they had successfully learned during training, ascertained by a post-training behavioural picture naming task. Participants had an average of 88% (SD = 12.0) accuracy on these participant-specific trained items in scanner and an average reaction time of 1054ms (SD = 203.5) in scanner. Participants also successfully learned semantic information about these successfully trained novel items, with an average gain of 83% (SD = 12.92). This level of variation in semantic knowledge of the new items ("Is it European?") demonstrates there was a continuum of

semantic consolidation between participants in the trained items. As would be expected, naming accuracy for already-known (pre- and post-training) items during scanning was high (M = 98%, SD = 2.2), with a mean reaction time of 1020ms (SD = 148.5). The naming latency for the newly learned and previously known items were not significantly different (t(19) = -.396, p = .696), indicating effectiveness of the training.

To explore the effect of semantic knowledge on word learning, correlations were performed between accuracy in the semantic judgement task ("Is it European?"), naming accuracy and subsequent maintenance of naming accuracy. There was a significant positive correlation between naming and semantic accuracy, with age at scan added as a controlled variable ($r(20) = .912$, p = .000). Additionally, there was a significant correlation between learning of the semantic cues and overall maintenance of learned, trained items ($r(20) = .66$, p = .002).

Whole brain results

The results of the whole brain analyses for the picture naming task are reported in Table 1, where three contrasts were created: 1) trained > untrained, 2) known > untrained and 3) trained > known. There were no significant clusters of activation for the opposing contrasts: untrained > trained, untrained > known and known > trained. There was a similar pattern of activation between the contrasts, where large bilateral language areas were identified. There was, however, greater and more extensive activation for the trained condition, including the hippocampus in both the trained > untrained and trained > known contrasts (Figure 2).

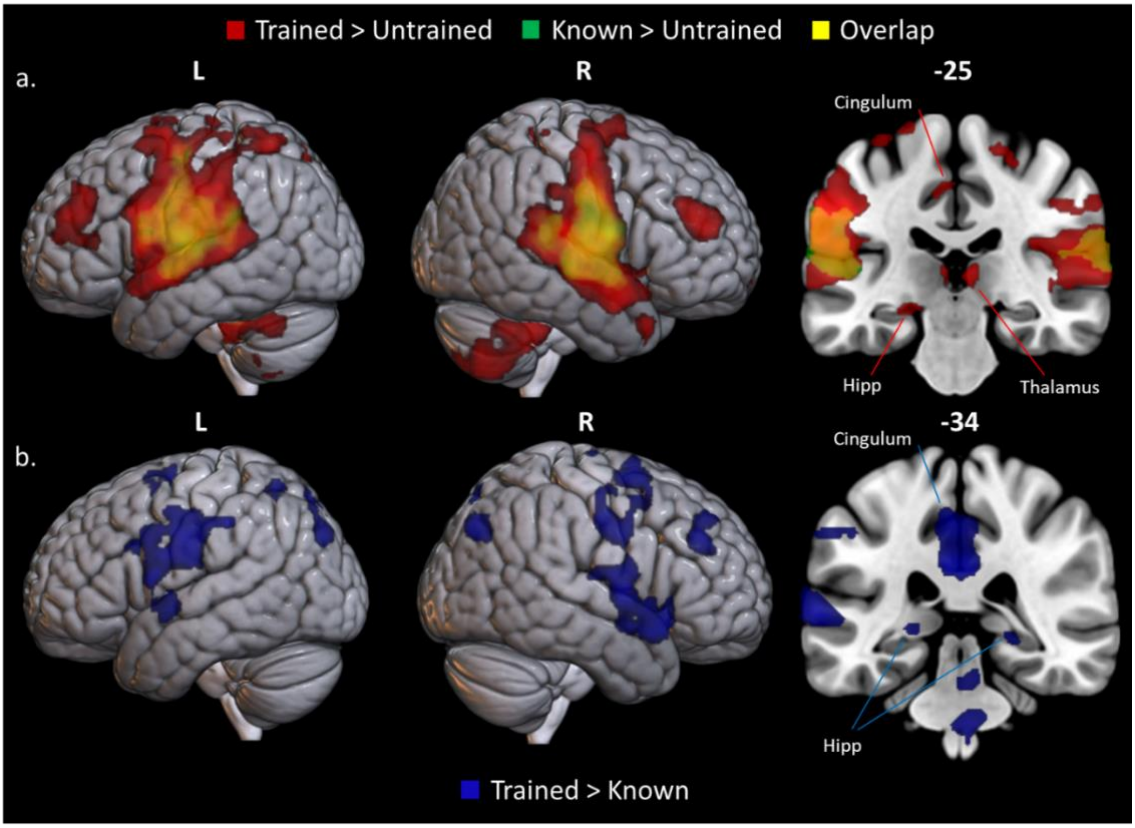


Figure 2. Whole brain BOLD activation of picture naming. A) trained minus untrained items (red) and known minus untrained items (green); yellow = overlap. B) trained minus known items (blue). Images thresholded at $p < 0.001$ voxel height, FWE-cluster corrected $p < 0.05$. L; left, R; right, Hipp; hippocampus, Thal; thalamus.

Table 1. Significant clusters of activation for naming of newly trained and already known items.

Contrast	Region of activation	Peak region	Cluster size	Peak MNI			*t*	*z*
				x	y	z		
Trained > untrained	L/R motor cortex, L/R frontal	L postcentral	42691	-56	-14	18	11.00	6.09
		R postcentral		60	-14	18	10.44	5.95
		L precentral		-48	0	22	10.11	5.92
	R MTL, R temporal lobe	R parahipp.	979	24	10	-22	5.90	4.41
		R TP		44	22	-32	5.68	4.29
		R parahipp.		22	14	-32	5.54	4.35
	L/R dorsal striatum, thalamus	L caudate	957	-12	-2	14	5.69	4.29
		R thalamus		4	-24	8	5.62	4.26
		L thalamus		-12	-8	14	5.32	4.08
	R MFG	R MFG	819	42	46	22	6.39	4.61
		R MFG		36	36	33	6.09	4.48
		R MFG		28	26	37	4.89	3.89
	L amygdala, orbitofrontal cortex	L amygdala	799	-16	-2	-12	5.97	4.44
		L SOG		-18	44	-16	5.71	4.30
		L MOG		-2	52	-10	5.42	4.12
Known > untrained	L motor, transverse gyrus, STG	L postcentral	4289	-60	-12	16	7.71	5.13
		L Heschl's		-48	-16	8	6.59	4.70
		L STG		-60	-20	8	6.50	4.66
	R temporoparietal	R postcentral	3078	60	2	16	8.46	5.39
		R postcentral		64	-10	18	6.84	4.80
		R STG		60	-18	2	6.27	4.56
	L cerebellum	L cerebellum	426	-20	-62	-22	5.69	4.30
	R cerebellum	R cerebellum	367	12	-62	-16	5.63	4.26

Note: Up to 3 strongest peaks listed per cluster, at $p < 0.001$ voxel height and $p < 0.05$ FWE cluster correction. Peak MNI = x, y, z. L; left, R; right, MTL; medial temporal lobe, TP; temporal pole, SMG; supramarginal gyrus, STG; superior temporal gyrus, IPL; inferior parietal lobe, IFG op; inferior frontal gyrus pars opercularis, IFG tri; inferior frontal gyrus pars triangularis, OFC; orbitofrontal cortex, MFG; middle frontal gyrus, AG; angular gyrus, parahipp; parahippocampal gyrus, SOG; superior orbital gyrus, MOG; middle orbital gyrus.

Table 2. Significant clusters of activation during picture naming in the trained > known contrast.

Contrast	Region of activation	Peak region	Cluster size	Peak MNI			*t*	*z*
				x	y	z		
Trained > known	L/R precuneus, cuneus, parahipp. gyrus, hippocampus	L cuneus	26037	-4	-74	26	9.02	5.56
		L calcarine		-12	-66	20	8.86	5.51
		L calcarine		26	-64	18	8.43	5.37
	L/R OFC, L insula, L IFG (p. tri),	L mid orbital	8968	-4	54	-6	9.86	5.80
		L insula		-34	22	2	7.96	5.22
		L IFG (p. tri)		-22	48	-16	7.52	5.06
	R insula, temporal pole, IFG (p. tri), MFG	R insula	6388	42	22	-2	8.41	5.37
		R MFG		36	38	28	6.56	4.69
		R MFG		26	16	46	6.45	4.64
	R AG	R AG	424	44	-78	40	5.60	4.25
		R AG		38	-52	40	4.81	3.85

Note: Up to 3 strongest peaks listed per cluster, at $p < 0.001$ voxel height and $p < 0.05$ FWE cluster correction. Peak MNI = x, y, z. L; left, R; OFC; orbitofrontal cortex, IFG; inferior frontal gyrus, MFG; middle frontal gyrus, AG; angular gyrus, parahipp; parahippocampal gyrus, SOG; superior orbital gyrus, MOG; middle orbital gyrus.

We did not find any significant clusters for the trained-untrained contrast in the semantic task; however, a significant cluster was identified for the untrained-trained contrast in the right anterior temporal lobe (rATL; peak MNI: 52, 4, -15) extending to the right superior temporal gyrus (STG; peak MNI: 56, 2, -12).

ROI analysis

To explore a core hypothesis arising from the CLS theory (a division of labour between MTL vs. cortical regions), behavioural data were correlated with activity in *a priori* ROIs related to episodic memory (bilateral hippocampi and left inferior parietal lobe; IPL) and semantic memory (left inferior frontal gyrus; IFG, left middle temporal gyrus; MTG and left anterior temporal lobe; ATL; Figure 3b). There were no significant correlations between semantic behavioural performance and *a priori* ROIs.

In the initial exploratory analysis, for the trained > untrained picture naming contrast (whereby participants named pictures of newly learned items, versus responding verbally to phase-scrambled images) we found a positive correlation between the left hippocampus and longer RTs (r = .519, p = .019; Figure 3c). Conversely, we observed inverse correlations in semantic areas located in IFG (IFG; r = -.528, p = .017; Figure 3c) and ATL, (r = -.611, p = .004; Figure 3c), where greater activation was related to quicker performance, suggesting they had deeper consolidation in the corresponding neocortical regions. There were no further correlations between trained > untrained BOLD and naming RTs. In the known > untrained contrast, there was a significant correlation between RT and left IFG BOLD activity (r = .509, p = .022; Figure 3c). There were no further significant correlations between known > untrained BOLD activity and RT, including the left hippocampus (r = -.106, p = .656) and left ATL (r = .001, p = .995; Figure 2c).

The key test of the CLS hypothesis is whether the trained > untrained behavioural correlations were significantly different from the known > untrained correlations, indicating differing neural networks for naming fully consolidated known items, versus less consolidated newly trained items (for a maximum of three weeks). The positive correlation of hippocampal activity in the trained > untrained contrast and RT, versus the weak negative correlation of hippocampal activity in the known > untrained contrast and RT, were significantly different using Fisher's r-to-z transformation (z = 1.846, p = .032, adjusted p = .032). In addition, the strong negative correlation of left ATL activity in the trained > untrained contrast and RT, was significantly different to the very weak positive correlation of ATL activity in the known > untrained contrast and RT (z = -2.25, p = 0.012, adjusted p = 0.018). Similarly, the negative correlation of left IFG activity in the trained > untrained contrast and RT was significantly different to the positive correlation of left IFG activity in the known > untrained contrast and RT (z = -3.348, p = .001, adjusted p = 0.001), Benjamini-Hochberg adjusted p values for multiple comparisons (Benjamini and Hochberg 1995), p = 0.05.

We also correlated in-scanner accuracy with BOLD activity for the trained > untrained contrast in each ROI. In the left hippocampus, individuals with greater activity showed poorer learning (r = -.456, p = .043; Figure 3a). Conversely, greater activity in the left ATL related to better accuracy (r = .450, p = .046; Figure 3a). Previously known items could be correctly named on three separate behavioural testing occasions, therefore,

there was high (M = 98%) accuracy on these items in the scanner, which does not provide variation for correlation with BOLD activity and therefore negates the ability to test the key hypotheses. These two correlations were significantly different to each other however, using Fisher's *r*-to-*z* transformation (z = -2.85, p = .004). All other correlations for trained > untrained accuracy, and known > untrained accuracy, with the a priori ROIs were not significant.

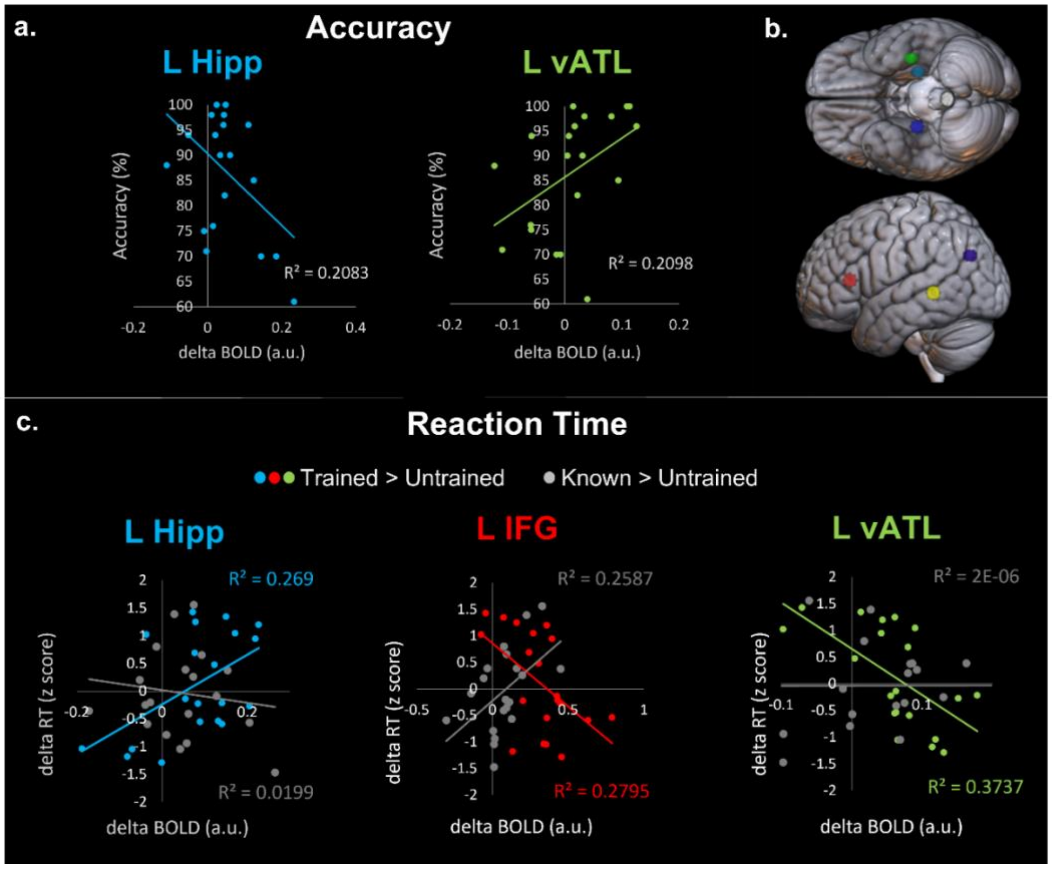


Figure 3. Correlations between functional activity and behavioural performance. A) Significant correlations of post-training percentage accuracy of trained items versus average BOLD for trained > untrained contrast. B) Spherical 6mm regions of interest: right hippocampus (navy; MNI: 28 -14 -15) left hippocampus (cyan; MNI: -28 -14 -15), left inferior parietal lobe (purple; IPL, MNI: -47 -64 3), left inferior frontal gyrus (red; IFG, MNI: -46 28 10), left anterior temporal lobe (green; vATL, MNI: -36 -15 -30), left middle temporal gyrus (yellow; MTG, MNI: -52 -42 0). C) Significant correlations between contrast estimates (coloured; trained > untrained, grey; known > untrained) and normalised in-scanner RT per participant per condition.

Maintenance data

Participants were re-tested on learned items five to six months post scanning, without interim training. Maintenance varied across participants, but overall participants named on average 73.9% (SD = 27.43) of learned words. To identify areas of BOLD activity which correlated with better or worse retention of trained items, percentage drop-off in naming performance over the maintenance period was added as a covariate of interest to the trained > untrained and known > untrained contrasts. In this covariate, higher values indicate worse retention of the trained words after the six-month maintenance period. With percentage drop off as a covariate of interest, over trained > untrained BOLD, we identified a cluster in the right hemispheric dorsolateral prefrontal cortex (rDLPFC, peak MNI: 38 8 46, Figure 4a). This indicates a correlation between more BOLD activity in the rDLPFC and greater trained item drop off (worse maintenance). There was no significant difference in the opposing direction (areas of BOLD correlating with better maintenance) or for the known-untrained contrast in either direction.

To explore the predictions from the CLS framework, we obtained a correlation between maintenance and the *a priori* ROIs during naming of trained words (Figure 4b). There was a significant positive correlation between left hippocampal activation and percentage drop off ($r(20) = .605$, $p = .005$), which suggests that individuals who were more reliant on hippocampal structures after learning, were less likely to retain the newly-learned vocabulary after a delay. There were no other significant correlations for trained > untrained or known > untrained contrasts.

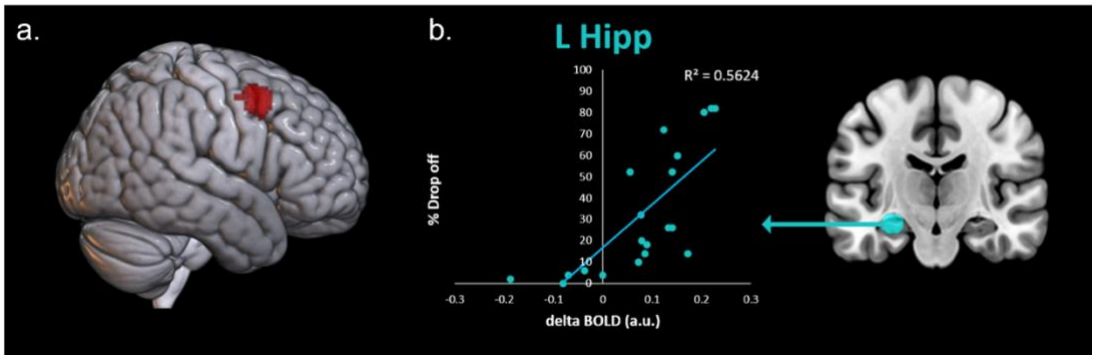


Figure 4. Correlations between maintenance of items and BOLD activity. a) Percentage trained item drop off as a covariate of interest in the trained-untrained contrast. Image thresholded at $p < 0.001$ voxel height and $p < 0.05$ FWE-cluster correction. b) Significant correlation between left hippocampal activity and percentage trained item drop off.

DISCUSSION

Vocabulary acquisition is a lifelong process for everyday life (e.g., 'coronavirus'), hobbies (e.g., 'thermocline') and careers (e.g., 'temporoparietal'). Reviving vocabulary is also key for individuals with language impairment after brain injury, stroke or dementia. This study evaluated the Complementary Learning Systems (CLS) framework (McClelland et al., 1995; McClelland et al., 2020) for the acquisition of novel real-world vocabulary in adulthood. At one time-point post-learning, a continuum of consolidation was demonstrated, with participants responding to completely unknown and untrained words, naming successfully trained words with varying levels of semantic knowledge, and naming previously known, well consolidated items.

The whole brain results indicate that new learning, in the trained condition, activates a combination of the typical language-semantic network, plus the hippocampal-episodic memory network. Whereas naming of well consolidated, previously known words, activates the cortical language-semantic network. The region of interest analyses demonstrated that activity in the left hippocampus during naming was associated with worse accuracy and slower reaction times, whereas activity in the language-semantic network (left inferior frontal gyrus, IFG, and left anterior temporal lobe, ATL) was associated with better accuracy and quicker reaction times. Additionally, the

maintenance results indicated that greater left hippocampal activity during newly trained naming was associated with greater drop off in item retention (i.e.,worse maintenance).

Complementary learning systems

The learning results described in this study fit within the CLS model. The CLS framework proposes a two-stage episodic-semantic account of learning; initial rapid hippocampal storage of new memories, followed typically by slower interleaved consolidation of new information alongside existing knowledge in the neocortex (McClelland, 2013; McClelland et al., 2020). In this study, at the whole brain level, in both the trained > untrained and known > untrained whole brain contrasts, activated clusters formed a typical motor/language network, including the inferior frontal gyrus (IFG). In addition, when recalling newly trained words but not when naming fully consolidated previously known words, we observed increased hippocampal activity (as observed in previous studies: Breitenstein et al. 2005, Davis et al. 2009) along with left inferior parietal lobe (IPL) activation. Our predictions were that naming newly trained words would rely on both episodic and semantic systems, whereas naming previously known, fully consolidated words would rely on the semantic-language systems only. These whole brain analyses support this notion. Region of interest analyses in combination with performance allowed us to explore this hypothesis in more detail.

In the episodic region of interest analyses for newly trained words, we found that left hippocampal activation was significantly associated with worse naming performance (less accuracy, longer RTs and less maintenance of trained words after 6 months). This effect was not found for the naming of previously known items, with only a non-significant weakly negative correlation. These two results were in line with our predictions, specifically, that individuals reliant upon the first MTL-episodic stage of the CLS would have worse performance for the newly-acquired vocabulary. It should be noted that we only found this effect in the left hippocampus, and not in the left inferior parietal lobe (IPL) or the right hippocampus. The previous literature has demonstrated a role of the left hippocampus in vocabulary acquisition (Breitenstein et al. 2005; Davis et al. 2009). As the language network is left dominant, it is logical that the episodic system supporting language acquisition is also left dominant. The left IPL has also been indicated in previous literature during word acquisition consolidation (Pohl et al. 2017). Although there was a significant cluster of IPL activation for the trained > untrained

contrast and not the known > untrained contrast, there were no significant correlations between IPL activation and behavioural performance. The functional organisation of the parietal lobe is complex, and although the region of interest was included as an episodic region based on previous literature (Wagner et al. 2005; Vilberg and Rugg 2008), various areas of the parietal lobe may be performing different functions, perhaps not aligning singularly with either the episodic or semantic network (Humphreys and Lambon Ralph 2015; Humphreys et al. 2020).

The neocortical areas activated by naming of newly learned items were typical of areas identified during speech production (Blank et al. 2002; Price 2012). We also identified two cortical regions associated with proficiency of naming learned items – the left anterior temporal lobe (ATL) and left inferior frontal gyrus (IFG). These areas are typically associated with semantic and language processing. The ventrolateral anterior temporal lobe is considered to be a trans-modal hub critical to semantic representation (Lambon Ralph et al., 2016). This proposal has strong, convergent support from multiple sources including semantic dementia patients (Acosta-Cabronero et al., 2011; Jefferies & Lambon Ralph, 2006; Patterson, Nestor, & Rogers, 2007; Warrington, 1975), fMRI (Binney, Embleton, Jefferies, Parker, & Lambon Ralph, 2010; Visser, Jefferies, Embleton, & Lambon Ralph, 2012), transcranial magnetic stimulation (Pobric et al. 2007, 2010), surface cortical electrode studies (Shimotake et al. 2015) and computational modelling (Rogers et al. 2004; Chen et al. 2017; Hoffman et al. 2018; Jackson et al. 2021). Sub regions of the LATL have been associated with picture naming and speech production specifically (Sanjuán et al. 2015). The IFG has been linked to speech production, amongst other processes, since Broca (1861) reported a patient with loss of articulation after destruction of the IFG and surrounding cortex. Despite debate as to the exact role of subregions of the IFG in speech production (Flinker et al. 2015) and semantic control (Whitney et al. 2012; Jefferies 2013; Noonan et al. 2013; Jackson 2021), the IFG is widely recognised to be important for articulation (Blank et al. 2002a; Hickok and Poeppel 2007; Lazar and Mohr 2011; Price 2012).

In these language-semantic regions of interest, we found an opposite pattern of results to those found in the episodic-hippocampal analyses. When naming newly trained items, more activity in the left ATL was associated with better accuracy and shorter RTs. In contrast, there was a non-significant weak positive correlation between ATL activation when naming previously known words. These results align with our

predictions that when individuals had better consolidated the new vocabulary (as indexed by better accuracy and shorter RTs) then this would be reflected in greater reliance upon the second neocortical stage of the CLS. This effect was also found in the left IFG, with activity during naming of newly trained items associated with quicker RTs. In addition, there was an opposite correlation of activity-behaviour when naming previously known items, whereby less activity in the left IFG was associated with quicker responses. This may reflect less neural effort for production of familiar vocabulary due to well-established phonological and articulatory representations (Blank et al. 2002; Price 2010, 2012) and/or fewer semantic control requirements. The fact that we observed (a) greater neocortical activity for the trained than known words and also (b) a negative correlation between activation and performance for the trained items, may well reflect the fact that not only should neocortical activation build up as the newly trained items are consolidated (and become independent of the MTL systems), but also that we know for established vocabulary from numerous language and semantic fMRI studies that there is more activation for less familiar/lower frequency words. Presumably, as proposed by many previous researchers, this reflects the fact that less frequent representations require more neural resources/longer processing times. Thus in the "life course" of new vocabulary one might expect an initial period in which the cortical activation builds up as the new vocabulary is cortically consolidated, but then with sufficient practice and use, the cortical representations should become more efficient and precise, thus be associated with decreasing cortical activation. This very pattern has been observed in implemented computational models of language (e.g., Chang and Lambon Ralph 2020) in which both initial vocabulary learning and relearning (after damage) is associated with an initial period of increasing unit activation and then a subsequent gradual reduction in unit activation as the underpinning (cortical) representations are more finely tuned.

It has previously been hypothesised that the CLS could apply in other domains (Davis and Gaskell 2009) and there are demonstrations in short-term pseudoword recognition (Cornelissen et al. 2004; Breitenstein et al. 2005; Mestres-Missé et al. 2007; Davis et al. 2009). Our findings complement and significantly extend these intra-learning investigations by exploring learning after full consolidation and maintenance of the new vocabulary. With three weeks training, the participants were able to name the items without cueing and make semantic decisions (i.e., more than exhibit above-chance

recognition performance). Taking this body of literature together, they clearly demonstrate that the hippocampal system is critical for new learning of artificial and native vocabulary learning, and that long-term consolidation reflects the gradual shift to long-term cortical representation and processing as predicted by the CLS model.

The speed of consolidation and reliance on the hippocampal-episodic network is now understood to be dependent on the strength of relationship between pre-existing knowledge and information to be learned (Kumaran et al., 2016; McClelland, 2013; McClelland et al., 2020). In this study, participants learned entirely new information (items, semantics and names). The item names are arbitrarily related to the object and their associated meaning; thus, this new knowledge is not systemically related to any pre-existing information. Therefore, the results obtained were as expected – it takes time to consolidate item names and, even after 2-3 weeks of learning, individuals remain reliant on a mixture of the hippocampal-episodic and semantic systems, rather than entirely on the cortical language-semantic system.

Methodological considerations

There were no significant clusters of activity for untrained > trained items. Previous literature has identified reductions of BOLD response related to word training (Nardo et al. 2017) and pseudowords versus word reading (Taylor et al. 2014). Activation can be interpreted as either engagement of the relevant systems or increased processing effort (Taylor et al. 2013). These areas of reduction can be interpreted as decreased effort associated with training. However, reductions may also signify responses to task difficulty, whereby items which are not trained are more difficult to respond to. In this study, items which were untrained were completely unknown to the participant, therefore the task is not inherently more difficult, as participants perform the same processes of viewing an image, thinking whether they know the name, then verbally responding.

Translational potential

This study is also potentially informative for aphasia therapy. The neural bases of successful speech and language therapy have been rarely explored, and those studies that have done so have yielded varying results (Abel et al. 2015; Nardo et al. 2017; Woodhead et al. 2017). The methods adopted in this study were deliberately designed to mimic those used to treat word-finding difficulties (where patients aim to re-establish

meaningful, native vocabulary through multiple learning sessions and vanishing phonemic cues (Abel et al. 2005), over several weeks (Dignam, Rodriguez, et al. 2016). By using the same paradigm, future studies can explore whether the neural correlates of word learning/re-learning in aphasia follows the same framework. The current results would seem to imply that therapy success will depend on: (i) the extent of damage to specific critical regions involved in the CLS framework; and (ii) damage to connectivity from the hippocampus to critical language regions. Furthermore, the majority of patients (especially those with middle cerebral artery stroke) tend to have intact hippocampus, which may be linked to the reason why patients experience initial success in learning, but long-term learning and maintenance (the goal of therapy) will relate to how well the therapy can induce relearning/stabilisation of neocortical representations. If these mechanisms hold in stroke aphasia, it could have important implications for intensity and dose of speech and language therapy provision.

Conclusions

The results of the study map the framework for word learning in the healthy older brain. In whole brain analyses, there was increased hippocampal activity when naming newly trained items, but not previously known, well-consolidated items. These results demonstrate the first stage of the CLS model, with initial hippocampal encoding. In addition, greater left hippocampal activity was associated with less accuracy, longer RTs and less maintenance of the newly trained words. When naming well-consolidated previously known words, there was no association between hippocampal activity and performance.

The second consolidation stage of the CLS proposes a gradual shift from reliance on the MTL-episodic network towards long-term neocortical consolidation. In line with this prediction, we found that when naming newly trained words, higher levels of left IFG and ATL activation were associated with better accuracy and shorter RTs. Crucially, the associations in each region of interest between BOLD activity and performance were significantly different between naming of previously known items and newly trained items. Overall, the results of this study provide evidence for both aspects of the CLS model in long-term, native word acquisition.

CHAPTER 2: TIME-LIMITED EPISODIC-LANGUAGE NETWORK CONNECTIVITY DURING NOVEL WORD LEARNING CONSOLIDATION

ABSTRACT

A prominent theory of knowledge acquisition is the Complementary Learning Systems model of memory, whereby a hippocampal system complements the neocortical memory system in a computational trade-off between learning episode specifics and regularities across episodes. The hippocampal memory system uses a high learning rate and sparse representations to store new knowledge from recent episodes quickly. In contrast, the neocortical system has a slow learning rate with deeper, overlapping representations to generalise across episodes. There is a gradual and changing division of labour across the two systems. However, it is unclear how these memory systems interact during varying stages of consolidation. We explored novel, native word learning in a healthy older population (n=20). Participants learned new words for three weeks, followed by a naming task during functional MRI acquisition. Naming already known, newly learned, and unknown/untreated words provided a snapshot of consolidation across word learning. Psychophysiological interaction analyses were performed to assess functional connectivity between a critical neocortical language region (left inferior frontal gyrus; IFG) and the rest of the brain at different stages of consolidation. When naming well-consolidated already known words, there was increased functional connectivity between left IFG and the picture naming network, regardless of behavioural performance. Conversely, when naming newly learned words, there was increased connectivity between the left IFG and the same naming network, plus areas involved in effortful episodic retrieval (left hippocampus and bilateral intraparietal sulcus) associated with decreased behavioural performance. These results suggest that there is a time-limited intermediate stage in which new knowledge is supported by both MTL-episodic and neocortical-language processing.

INTRODUCTION

Spoken languages are constantly evolving with new words regularly introduced from technological innovation, foreign assimilations, and informal terms. Additionally, we are exposed to new vocabulary across the lifespan reflective of our continuing experience. New professions and hobbies require specific novel terms and jargon. Thus, throughout the lifespan, it is imperative that individuals can acquire new words for quality of life (Gabriel and Bowling 2004). A prominent theory of knowledge acquisition is the Complementary Learning Systems hypothesis (CLS; Kumaran et al., 2016; Marr, 1971; McClelland, 2013; McClelland et al., 2020, 1995). In the CLS model, the neocortex implements a gradual, interleaved learning strategy. This slow strategy allows for only subtle changes on the basis experience and thus, over-generalisation of concepts, or catastrophic interference does not occur. However, humans can learn knowledge very quickly. Thus, rapid but sparse learning supported by the medial temporal lobes (MTL) allows for quick acquisition of new knowledge without catastrophically disrupting the structured system of knowledge gradually built up through experience (Gaskell and Dumay 2003). Thus, both complementary systems are needed to perform the complex requirements of both the rapid acquisition of representations and gradual integration of new concepts with pre-existing knowledge (McClelland et al. 1995). The CLS model has been theorised to apply to vocabulary acquisition (Davis and Gaskell 2009). This implies that processing supporting the consolidation of new vocabulary shifts gradually from MTL-episodic to semantic-language related regions.

Yet, how these complementary processes interact remains unclear. Traditionally, based on neuropsychological evidence, the episodic and semantic systems have been considered functionally distinct (Tulving 1972, 1985) based upon neuropsychological evidence. Patients with MTL damage have severe anterograde and retrograde deficits, with relatively spared semantic memory (Manns et al. 2003; Bayley et al. 2006; Rosenbaum et al. 2008). Whereas patients with semantic dementia (SD), the progressive deterioration of anterior temporal lobe, exhibit gradual semantic deficits with relatively unimpaired episodic memory (Hodges et al. 1992). Despite this apparent functional distinctiveness, the need for complementary memory systems (McClelland et al. 1995)

highlights the critical nature of their interactivity via a gradual shift in the division of labour between episodic and language-semantic systems.

This gradual shift in the division of labour between episodic and language-semantic systems is evident from short term, online word learning studies. Breitenstein et al. (2005) used an associative learning paradigm of pseudoword-picture pairings. Over the one hour incidental learning period, there was a gradual decrease of hippocampal (and left fusiform gyrus) activity associated with increases in associative learning. Davis et al. (2009) presented participants with three sets of novel pseudowords, of which two have received training a day before scanning. Presentation of entirely novel pseudowords was associated with increased hippocampal activation, which decreased with subsequent presentation repetitions. In contrast, presentation of items which were subjected to overnight consolidation was associated with increases in activation for language areas such as left inferior frontal regions. Although these studies demonstrate shifts between complementary systems during initial encoding of novel representations, there is evidence that consolidatory processes can occur over a more extended period of weeks or months. In Gore et al., (2021) the data of which are reanalysed in this study, participants learned up to 50 novel, native words over three weeks. Post-training, participants named newly learned items, already known and thus well-consolidated items and unknown/untrained items. When retrieving newly learned item names, there were areas of increased activity in both episodic and language-semantic regions. Therefore, there was still evidence of a division of labour between memory systems after weeks of training.

Exploring simple levels of activation provides evidence of the CLS model in word learning but can depict only the average level of engagement across different regions within distributed systems. This does not necessarily explore how these systems interact to delineate a gradual shift in the division of labour during consolidation. Functional connectivity denotes areas of activity correlation between regions (Friston et al. 1994) and thus reflects how larger scale neural systems are functionally coupled to achieve cognitive tasks. Although not explicitly tested in the previous literature, the results of these online learning studies hint at gradual connectivity changes across the learning time course. To further explore these suggested functional connectivity changes, long-term functional connectivity studies of newly learned word recall are needed, allowing for consolidation time including sleep periods.

Although the discussed previous literature suggests episodic-language connectivity during recall of newly learned words, there is debate as to the time course of this hippocampal-neocortical interaction in memory. The CLS model predicts a time-limited role of hippocampal support, supported by neuropsychological and animal models of medial temporal lobe injury resulting in amnesia (cf. Squire and Alvarez, 1995). Following Ribot's law (Ribot 1881), amnesia following brain injury is often temporally-graded, whereby recently acquired knowledge is more likely to be lost than remote memories. Therefore, in standard models of consolidation, the hippocampus is no longer required for semantic representations after some time. Conversely, a growing body of literature purports a permanent role of the hippocampus in semantic memory (Nadel and Moscovitch 1997; Sekeres et al. 2018; Duff et al. 2020) during both replay and task performance. Intracranial recordings suggest a relationship between hippocampal theta power and tasks involving semantic processing, such as picture naming and verbal free-recall (Hamamé et al. 2014; Solomon et al. 2019). However, without clear baselines, it is unclear whether this hippocampal activity is associated with semantic processing or the episodic encoding of the experience of each trial. Episodic memory systems are automatically engaged when presented with a perceptual event, such as within task trials. Thus an active baseline is needed to account for this activity (Martin et al. 1997; Gilmore et al. 2021). Therefore, the question remains whether the role of hippocampal activity in language-semantic knowledge acquisition is time-limited.

This study aimed to explore whether there is functional connectivity between the episodic and semantic-language networks during naming of newly learned items. If so, is this connectivity time-limited to the intermediary stages of consolidation, or is there evidence of a permanent hippocampal role during picture naming? To investigate these questions, we performed task-based functional connectivity analyses (psychophysiological interaction; PPI; Friston et al., 1997; O'Reilly et al., 2012) on data from a previous study of healthy older age word acquisition (Gore et al., 2021; Chapter 1). Functional imaging data included picture naming across a spectrum of word consolidation stages, from well-consolidated already known items to an intermediary step of newly learned words and finally unconsolidated and unknown items.

To determine a seed region for PPI, a joint analysis of the univariate naming contrasts during picture naming minus the baseline (newly learned or already known > untrained)

was performed to identify brain regions with increased activation during naming in both stages of consolidation. There were multiple clusters of increased activation when naming was compared to the baseline, including the left inferior frontal gyrus (IFG). The left IFG has been linked to speech production, amongst other processes, since Broca (1861) reported a patient with a loss of speech production after damage to the IFG and surrounding cortex. Large scale reviews of the previous literature indicate left IFG involvement in naming (Blank et al., 2002; Price, 2011; Price et al., 1996; Price, 2010). A critical level of damage to this region is also associated with decreased word learning performance (see Chapter 5). In the data used in this study, activation of the left IFG was associated with better behavioural performance when naming newly learned words (Gore et al., 2021; Chapter 1). Thus, the peak coordinate of the left IFG cluster of increased activation was used as the PPI seed region.

We predicted that there would be an intermediary stage between initial rapid encoding and long-term retention. Consequently, we expected increased functional connectivity between the essential IFG language-semantic PPI seed and episodic regions when naming newly learned words, in contrast to no difference in connectivity when naming words that were already known and therefore well consolidated. We expected that reduced naming performance would be indicative of lesser consolidation. Consequently, we predicted increased functional connectivity between language-semantic and episodic regions would be associated with worse behavioural naming performance.

METHODS

This study involved functional connectivity analyses of previously reported univariate data (for further information: Gore et al., 2021). A summation of the associated methods follows.

Participants

Participants were recruited from the North West of England (twelve females, eight males, aged 46-76 years, M = 63.90, SD = 8.82). All participants were right-handed with no history of dyslexia, language deficits or neurological disease, including dementia (assessed with the Addenbrooke's Cognitive Examination Revised (ACE-R;

cut-off score of 88, Mioshi et al., 2006). Participants had no contraindications to MRI and normal/normal-to-corrected vision. Participants gave informed consent before participation. Ethics were approved by a local National Health Service (NHS) ethics committee.

Procedure

Picture naming stimuli were high quality, coloured images of real-world nouns. Items ranged from high frequency to low frequency, selected from the British National Corpus (BNC Consortium, 2007). Initially, participants were tested on picture naming of all items. Items which could be named correctly were added to the already known stimuli set. Items which could not be named were randomly split into two sets: to-be-trained and untrained items. Sets were matched for psycholinguistic variables and counterbalanced. Phase-scrambled images of items from the other three conditions were used for the baseline condition. Participants received three weeks training on to-be-trained items, henceforth termed newly learned items.

Two types of well-studied speech and language therapy approaches were used for training. The first two weeks involved choice of cue training (Nickels 2002; Abel et al. 2005; Pohl et al. 2017). Initially, 10 items were trained. After a participant reached 70% accuracy, an additional 10 items were added up to a total of 50 items. The third week consisted of repeated increasingly speeded presentation (RISP; Conroy et al., 2018) on the same training items. When participants reached a speeded success rate of 70%, response times were incrementally reduced from 1.8s to 1.4s to 1s. Post-training, participants were re-tested on all trained items. Only successfully trained items were added to participant-specific newly learned fMRI item sets.

Neuroimaging acquisition

Structural and functional scans were acquired using a Phillips Achieva 3.0 Tesla system with 32 channel SENSE coil with 2.5 sense factor. Verbal responses were recorded with a FOMRI microphone and noise-cancelling Mk II+ headphones (MR Confon) were worn inside the scanner. 260-slice structural images were acquired: TR = 8.4ms, TE = 3.9ms, flip angle = 8 degrees, FOV = 240 x 191mm, resolution matrix = 256 x 206, voxel size = 0.9 x 1.7 x 0.9mm. A triple echo planar imaging technique was used for functional image acquisition. Acquiring three echoes (12, 32 and 48ms) allowed for less

signal loss in areas of magnetic susceptibility in key episodic and language-semantic regions such as the ventral temporal lobes and medial frontal cortices (Poser et al. 2006; Halai et al. 2014, 2015b). Additionally, a 45-degree upward tilt from the AC-PC line was used to reduce ghosting artefacts in the temporal lobes. 31 whole brain slices were acquired with the following parameters: TR = 2.5s, TE = 12, 30 and 48ms, flip angle = 85 degrees, FOV = 240 x 240mm, resolution matrix = 80 x 80, and voxel size = 3.75 x 3.75 x 4mm.

A picture naming task comprised a block design with four conditions; already-known, newly learned, untrained and baseline. Participants overtly named or responded 'don't know' to pictures in all conditions. Responses were recorded using a fibre optic fMRI microphone (FOMRI; Optoacoustics). Each 1900ms trial consisted of a fixation cross for 700ms, followed immediately by the target image in the middle of a white screen for 1200ms. Each block lasted 9.5s containing five trials. Eight jittered rest blocks per run were included with an average length of 9.5s. With 32 task blocks and eight rest blocks per run, the total run time was 6 minutes and 33 seconds. In separate runs participants performed an irrelevant task for the purposes of a separate study.

Neuroimaging pre-processing and analysis

T1 data was pre-processed using the FMRIB Software Library (FSL, version 6.0.0; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/, Woolrich et al., 2009). The skull was stripped from the structural T1 images (BET; Smith, 2002) including sub-cortical segmentation. A bias-field correction was applied using FSL's anatomy pipeline (FAST; Zhang, Brady, & Smith, 2001). Registration to standard space was performed in FSL with FLIRT and FNIRT (Woolrich et al. 2009; Patenaude et al. 2011) and segmentation with FAST (Zhang et al. 2001). Despiking and slice time correction was applied to the functional data in the AFNI neuroimaging suite (v19.2.10; Cox, 1996; Cox & Hyde, 1997; 3dDespike; 3dTshift). For each functional echo, combined normalisation, co-registration and motion correction parameter sets were applied in FSL. Functional data were optimally combined, taking a weighted summation of the three echoes, using an exponential T2* weighting approach (Posse et al. 1999) and regression analysis. Functional runs were also combined and denoised using multi-echo independent component analysis (ME-ICA; Kundu et al., 2013; Kundu, Inati, Evans, Luh, & Bandettini, 2012) using the tool meica.py (v3.2) in AFNI (Cox, 1996; Cox & Hyde,

1997). The denoised timeseries were normalised to standard space using FNIRT warps, then smoothed.

PPI

Psychophysiological interaction (PPI) analyses are used to identify task-based functional connectivity changes (Friston et al., 1997; O'Reilly et al., 2012). These analyses reveal how task conditions modulate the relationship the activity of a seed region and other brain areas (over and above their baseline connectivity and individual levels of activation in response to the task). Thus, PPI analyses establish which voxels demonstrate differential connectivity with the seed region as a function of task condition.

The aim of the PPI analyses was to explore functional connectivity between episodic and language-semantic regions during picture naming. To determine a seed region for PPI, a joint analysis of the naming contrasts over baseline (newly learned + already known > untrained) was performed to identify brain regions with increased activation during naming in both stages of consolidation. There were two large bilateral clusters of activation across motor cortices, spanning to the inferior frontal gyrus (IFG), superior temporal lobe, superior parietal lobe and inferior occipital lobe (Figure 1a). The left IFG is a key region involved in speech production and naming, (Blank, 2002; Price, 2012, 2010) and critical levels of damage to this region is associated with reduced word learning performance (see Chapter 4). Increased activity over baseline in the left IFG was also associated with better behavioural performance in this data when naming newly learned words (Gore et al., 2021; Chapter 1). Thus, the peak coordinate of the left IFG cluster was used as a 6mm spherical PPI seed (MNI: -38 28 16).

The time course of the left IFG seed region was extracted. PPI analyses were performed to identify brain regions which showed differential connectivity with this left IFG region for the newly learned > untrained and already known > untrained conditions. This required constructing a first level general linear model, using the PPI time course (a regressor describing the interaction between the time course of the region and the time course of each task) and including the physiological and psychological time courses as covariates. To further test the hypothesis that increased task demands require increased language-semantic and episodic network functional connectivity, a parametric

modulator of reaction time (RT) was added to the PPI GLM analyses for both naming conditions (newly learned and already known).

These analyses were performed using statistical parametric mapping (SPM12) software (Wellcome Trust Centre for Neuroimaging) and MarsBaR. Reaction times (RT) were calculated from stimulus onset. Inverse efficiency scores (IES) incorporate both accuracy and RT to gain an overall impression of behavioural performance and may be less biased by an individual's response strategy than RT or accuracy alone (Townsend and Ashby 1978). IES were calculated by dividing RT by one minus the proportion of correct responses. To correct for multiple comparisons when performing multiple correlations between behavioural performance and functional connectivity, *p* values were Benjamini-Hochberg corrected with a critical level of .05 (Benjamini and Hochberg 1995).

RESULTS

To explore the hypothesis that there is increased connectivity between language-semantic and episodic regions during novel word learning consolidation, but not during naming of well consolidated already known items, a PPI analysis was performed with left IFG connectivity in the newly learned > untrained contrast and already known > untrained contrasts respectively. The left IFG peak seed showed greater functional connectivity during naming of already known words than verbal responses to untrained items with early visual and ventral occipitotemporal (vOT) regions (Figure 1b). In the newly learned > untrained contrast, there was increased connectivity with the left IFG peak seed in this picture naming network. However, there were also areas of significant increases in connectivity with left IFG in bilateral intraparietal sulcus (IPS) and left pre-motor regions (Figure 1b). There were no areas of decreased functional connectivity with the left IFG peak seed in these contrasts.

To investigate whether reduced naming performance was associated with increased functional connectivity between the left IFG and any brain region, IES was added as a parametric modulator to the whole-brain PPI analysis. IES incorporates accuracy and reaction times to give an overall impression of behavioural performance. Areas of increased functional connectivity with the left IFG peak seed that correlated with longer

RTs/reduced accuracy included areas of the MTL including the hippocampus, bilateral intraparietal sulcus (IPS), right angular gyrus and areas of the cerebellum (Figure 1c). There were no significant correlated regions of functional connectivity associated with faster RTs/increased accuracy. There were also no areas of significant functional connectivity changes associated with behavioural performance in the already known > untrained contrast.

To further explore the relationship between behavioural performance and inter-regional connectivity and indicate directionality, areas of significant correlation between functional connectivity and behaviour were used as ROIs. Higher IES scores indicate slower RTs or decreased accuracy. In the newly learned > untrained contrast, left IFG to left hippocampus functional connectivity significantly positively correlated with IES (r = .755, p = .0001, corrected p = .0006, Figure 1c). There was no significant correlation between left IFG to left hippocampus connectivity and IES in the already known > untrained contrast (r = .0007, p = .913, corrected p = .913, Figure 1c). The same pattern was demonstrated in the left IPS cluster of significant connectivity. In the newly learned > untrained contrast, left IFG to left IPS functional connectivity significantly positively correlated with IES scores (r = 0.511, p = .012, corrected p = .036). In contrast, there was no significant correlation between left IFG to left IPS connectivity and IES scores in the already known > untrained contrast (r = .328, p = .159, corrected p = .239). There was also a significant positive correlation in left IFG – right AG connectivity and IES in the newly learned > untrained contrast, (r = 0.507, p = .0452) and no significant correlation between the same variables in the already known > untrained contrast, (r = .076, p = .750, corrected p = 0.9).

Table 1. Areas of significant functional connectivity with the left IFG during picture naming.

Contrast	Region of activation	Peak region	Cluster size	Peak MNI			*t*	*z*
				x	y	z		
Newly learned > untrained	L/R posterior OTC, L/R precuneus	R fusiform	18859	46	-62	-18	7.80	5.16
		R cerebellum		6	-78	-36	6.60	4.70
		L lingual		-20	-88	-16	6.51	4.66
	L IFG, L motor	L precentral	1303	-46	-6	48	5.99	4.43
		L IFG tri.		-52	18	30	5.19	4.04
		L precentral		-46	4	24	4.08	3.42
Already known > untrained	L/R OTC	R occipital	1997	30	-96	6	5.80	4.35
		R MTG		52	-68	2	5.29	4.10
		R fusiform		44	-60	-16	5.21	4.06
	L OTC, L cerebellum	L fusiform	777	-38	-78	-12	4.78	3.83
		L cerebellum		-36	-70	-18	4.47	3.65
		L fusiform		-38	-48	-22	4.32	3.56
IES: Newly learned > untrained	L MTL	L hippocampus	896	-20	-32	-10	5.24	4.03
	L/R parietal	L IPS	1362	-30	-74	52	4.82	3.81
		R IPS		20	-72	28	4.57	3.68
		L precuneus		2	-72	58	4.48	3.63
	R AG	R AG	664	50	-58	28	4.81	3.81
	Cerebellum	L cerebellum	821	-8	-88	-20	4.64	3.72

Note: Three strongest peaks per cluster listed. Clusters significant at $p < .005$ voxel level, $p < .05$ FWE cluster correction. L, left; R, right; OTC, occipitotemporal cortex; IPS, intraparietal sulcus; IFG tri., inferior frontal gyrus (pars triangularis); IFG op., inferior frontal gyrus (pars opercularis); ITG, inferior temporal gyrus; MTG, middle temporal gyrus.

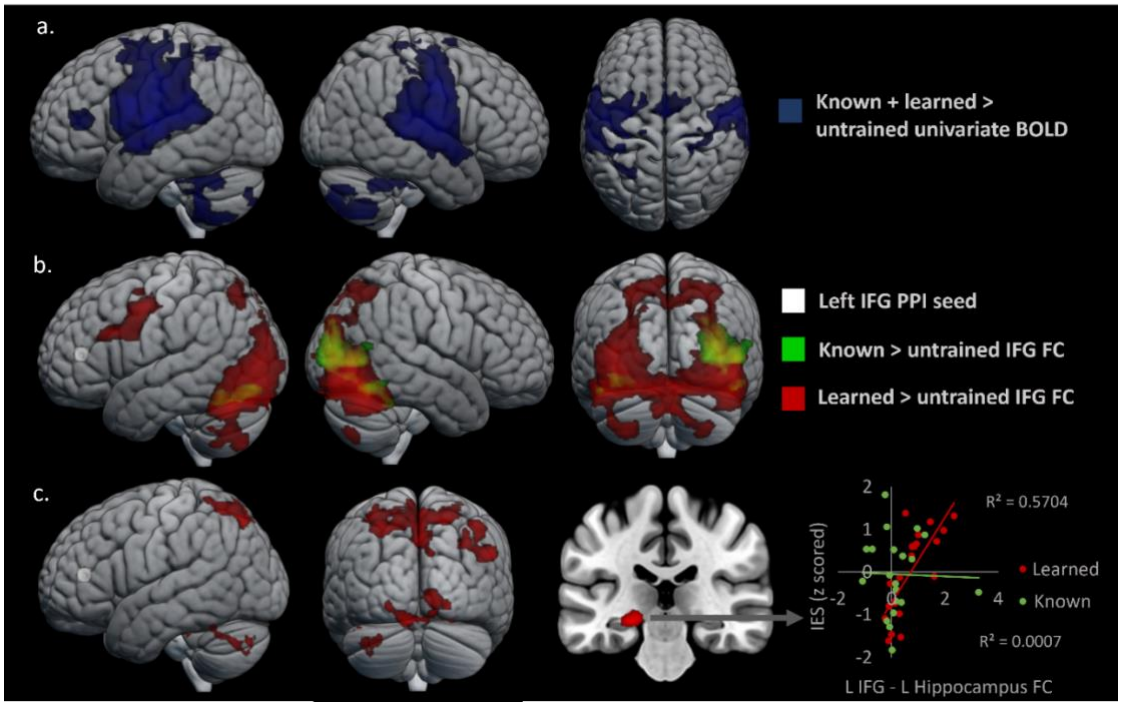


Figure 1. Regions of increased functional connectivity. A) Areas of increased BOLD activity in the newly learned + already known > untrained contrast in the picture naming task. Univariate results thresholded at $p < .001$ voxel level and FWE-corrected critical cluster level of .05. B) Areas of greater functional connectivity (FC) with the left IFG PPI seed (white) for the newly learned > untrained (red) and already known > untrained (green) contrasts. Areas of overlap are indicated in yellow. C) Areas of greater connectivity to the left IFG seed (white) modulated by inverse efficiency scores (IES) in the newly learned > untrained contrast (red). Scatterplot of left IFG functional connectivity with left hippocampus and IES in both the newly learned > untrained (red) and already known > untrained (green) contrasts. PPI results thresholded at $p < .005$ voxel level and FWE-corrected critical cluster level of $p < .05$.

DISCUSSION

The overarching aim of this study was to begin to disentangle the interactivity of episodic and language-semantic systems across a continuum of word learning consolidation stages. Functional connectivity with the left IFG for the naming of established vocabulary revealed part of the network that is commonly found for picture naming and speech production (ventral occipitotemporal (vOT) and early visual

regions). When naming newly learned items, there was increased functional connectivity with the left IFG in the same areas of the naming network and left motor/premotor regions, plus, key episodic retrieval areas associated with effortful retrieval (bilateral posterior IPS). To explore whether there were differences in functional connectivity dependent on behavioural performance, and thus level of consolidation within the newly learned items, IES were considered. Increased left IFG functional connectivity during naming of newly learned items with the bilateral IPS, right AG and left MTL was associated with slower RTs/reduced accuracy in naming newly learned items, and not associated with naming of already known items. Thus, a gradual change in the division of labour was demonstrated, with less well consolidated items driving inter-regional functional connectivity.

Whole brain functional connectivity with left IFG

The areas associated with increased functional connectivity with the left IFG in naming of already known items (early visual areas and vOT cortex) were typical of those found in multiple studies of object naming (Bookheimer et al. 1995; Price et al. 1996; Moore and Price 1999; Murtha et al. 1999; Price 2012). To name an object, the stimuli must be visually processed to identify the various features relevant to stored representations of items (Humphreys et al. 1999). Early visual and vOT are involved in the integration of visual information with higher-level processing (Price and Devlin 2003). As the left IFG (the PPI seed region) is consistently activated in picture naming studies and phonological tasks, these connectivity results in naming known items can be considered typical integration of visual information with higher-level phonological processing during object naming. There was also significant connectivity between these same typical visual areas and the left IFG when naming newly learned words, but with the addition of regions associated with episodic processing (bilateral IPS). Posterior IPS activation is related to effortful episodic retrieval (Wagner et al. 2005). Therefore, in the intermediary stage of word learning consolidation only, there was increased functional connectivity between the typical naming network and episodic processing areas.

Functional connectivity and behavioural performance

To explore the continuum of consolidation within the intermediary stage of newly learned words, the behavioural measures of RT and inverse efficiency scores (IES) were considered. The same episodic regions were also associated with the behavioural

measure of IES. This means that the less well consolidated items were the ones that drove the significant increases in inter-regional connectivity. Thus, within the intermediary stage of consolidation, there was a continuum whereby level of connectivity between episodic and language regions was mediated by level of consolidation. As participants consolidated picture-name representations, connectivity between the typical naming network and areas associated with effortful episodic retrieval reduced. This result fits within the broader episodic literature, which indicates that posterior IPS are implicated in effortful episodic retrieval (Wagner et al. 2005). Notably, there were no areas of increased connectivity with the left IFG associated with behavioural performance when naming already-known items.

How long does it take for the division of labour to shift?

The results of this study add to the debate of whether the role of the hippocampus in semantic processing is permanent (Nadel and Moscovitch 1997; Sekeres et al. 2018), or, whether there is a time-limited role during systems consolidation (McClelland et al. 1995). Previous research has demonstrated episodic network activity during consolidated word recall (Hamamé et al. 2014; Solomon et al. 2019), but the cause of this activity is debated. To retain and learn knowledge, episodic memory systems must continually update with new information (Martin et al. 1997). It may be the case that when items are displayed in the scanner, the episode of the item trial itself is re-acquired, which would account for any episodic activity. Therefore, it is vital to have an active baseline task that may stimulate episodic encoding of information. The baseline of this study involved viewing new, phase-scrambled images and responding verbally. Thus, any episodic activation (or coactivation) should be above and beyond the episodic encoding of a trial.

Two elements of these results provide evidence of this time-limited role. Firstly, as there was no functional connectivity between episodic regions and the left IFG during naming of well-consolidated items, the results suggest that hippocampal involvement in lexical retrieval is time limited. However, the already known items in this study are likely to have been learned decades earlier by participants as they have a low age of acquisition. As such, the lack of episodic connectivity in the naming of known words only provides information as to an end point of hippocampal involvement, once words are sufficiently consolidated, but not a specific timescale of involvement. Length of

consolidation can depend on the relatedness of new information to previously learned information. Some types of information still rely on episodic contributions for a long time (McClelland et al. 2020). In this study, participants learned entirely new information with limited lexical-semantic association, unrelated to pre-existing information. Therefore, a longer intermediary period of three or more weeks is as expected.

Exploring the association between network connectivity during naming and RT, allowed for a deeper inspection of the time periods involved in the gradual shift of divisions of labour. RTs decrease in alignment with increasing consolidation status (Davis et al. 2009). Individuals with longer RTs for newly trained items (controlled for general naming speed) demonstrated more connectivity between episodic and language networks. Thus, within the three weeks of training there were differential levels of consolidation and associated differences in interregional connectivity. To further explore the timescale of this shifting division of labour a longer period of training is needed, with scanning taking place pre-training, during training, and at periods post-training.

Conclusions

In conclusion, we aimed to explore whether there is episodic-semantic connectivity during the active recall of newly learned words and whether this episodic involvement is time-limited. We used newly learned native words as an intermediary between already-known, well consolidated words and completely unknown items. From a left IFG functional connectivity seed, we observed increased connectivity between areas associated with language and episodic processing during naming of newly learned items only. This episodic-language connectivity was not demonstrated when naming already known items. Additionally, connectivity between episodic and neocortical language regions was associated with worse performance, indicating items were less well consolidated. These results provide evidence for an intermediary stage during word learning of complementary episodic and neocortical-language processing that is time-limited, indexed by consolidation.

CHAPTER 3: BILATERAL SYSTEMS IN ADULT LANGUAGE LEARNING: A VARIABLE NEURO-DISPLACEMENT ACCOUNT

ABSTRACT

Although the functional neural correlates of novel language learning have received some attention, there remains relatively little known concerning the role of variation in structure amongst healthy participants. In this study, healthy older adults (n = 20, age = 46-76 years, mean age 63.90) learned 50 new real-world native English nouns. Voxel-based correlational methodology (VBCM) identified right-sided neural correlates of increased word learning ability. In a region of interest analysis, right inferior frontal gyrus structural intensity measures correlated with better behavioural performance and less activation of episodic regions for naming of newly trained items. These associations were moderated by structural intensity of the left inferior frontal gyrus. Individuals with higher structural intensities in the dominant left hemispheric language regions did not demonstrate right-sided functional or structural associations. Only when there was decreased structural intensity in critical frontal regions was the right hemisphere required structurally. This relationship is an example of a mechanistic process determining lateralisation of language learning correlates dubbed "variable neuro-displacement", a principle that may be applicable across a number of different populations.

INTRODUCTION

Although the functional neural correlates of novel language learning have received some attention, relatively little is known concerning the role of variation in structure amongst participants. Language is typically viewed as a unilateral, left hemispheric dominated process. This notion is supported by neuropsychological evidence of chronic aphasia occurring from left sided damage, not right hemisphere damage (Lichtheim 1885). However, recent research suggests that right hemispheric lesions can result in language deficits (Gajardo-Vidal et al. 2018). Additionally, in healthy participants, several types of language tasks elicit bilateral activation, including repetition, comprehension and production (Blank et al. 2002a; Bozic et al. 2010; Poeppel 2014).

The mechanisms underlying recovery from post-stroke aphasia have been proposed to fit within a set of mechanistic ideas termed "variable neuro-displacement" (Stefaniak et al. 2020), a term borrowed from engineering (Manring and Johnson 1996). This set of ideas comes from a fundamental principle of engineering that well-engineered systems are resilient to functional stresses whilst balancing performance and energy costs. Aphasia recovery can be conceptualised as engagement of degenerate networks or spare capacity within or between networks via variable neuro-displacement. Degeneracy refers to networks or regions that were not premorbidly involved in language processing but are engaged for language tasks after damage. These networks may include undamaged right hemispheric language homologues.

Additionally, there may be existing spare capacity within or between networks. Under standard levels of performance demand, this spare capacity may be downregulated to reduce energy costs. When task demand increases or after network efficiency is reduced via damage, activity in these areas may be upregulated (Stefaniak et al. 2020). Upregulation could involve language and domain-general executive processing networks. This principle may explain findings that increased activation in perilesional and language homologue regions correlates with patient performance post-stroke (Heiss et al. 1999; Saur et al. 2006; van Oers et al. 2010; Szaflarski et al. 2013).

Although the idea of variable neuro-displacement aligns with research on post-stroke aphasia recovery (Chang and Lambon Ralph 2020; Stefaniak et al. 2020, 2021), it is unclear whether these ideas can be applied beyond aphasia as a general principle. Many situations induce variable performance demands in neurotypical populations. Firstly,

some tasks are inherently more difficult than others, and the difficulty of a given task may vary based on the specific task demands. For example, degrading visual stimuli makes object recognition harder. Similarly, more difficult visual processing commonly occurs in real-life, such as when driving in fog. Increased metabolic energy consumption and spare capacity usage within networks may be biologically beneficial to allow sufficient resources to get home safely even in this more demanding context. In the context of word learning, naming of familiar items, which were acquired a long time ago and therefore are well-consolidated, is relatively easier than naming newly learned items. Secondly, tasks can become more difficult over the lifespan due to typical ageing processes. There is documented cognitive decline due to ageing in most domains due to extensive tissue loss of grey (Good et al. 2001; Resnick et al. 2003; Raz et al. 2005) and white matter (Sowell et al. 2003; Giorgio et al. 2010). Thirdly, there may simply be individual differences in how well participants perform a process, and therefore how easy they find a particular task. These individual differences may be innate or related to prior experience.

Previous structural studies demonstrate that measures of novel syllable production (Golestani and Pallier 2007), second language proficiency (Mechelli et al. 2004) and adolescent knowledge of vocabulary (Lee et al. 2007) correlate with structural intensities in the supramarginal gyrus (SMG) within the inferior parietal lobe (IPL). Functional studies indicate that word learning consistently activates the IPL (Breitenstein et al., 2005; Cornelissen et al., 2004; Gore et al., 2021; Chapter 1). As such, grey matter density of the IPL has been suggested to be a neural marker for vocabulary acquisition success regardless of age (Green et al., 2007). It is not yet known whether these correlations between structural intensities and word learning are limited to the left hemisphere.

This study investigated the structural correlates of word learning ability to determine whether the principles of variable neuro-displacement can usefully be applied to a neurotypical population. Older participants were recruited for comparability to word learning or re-learning in post-stroke aphasia to improve understanding of the premorbid state of the language system in aphasic patients. 20 older participants learned 50 novel items and associated semantic information, as reported in Gore et al. (2021; Chapter 1). These data were further probed to explore the following questions. Are areas of increased structural intensity associated with better learning of novel words

limited to the left hemisphere? If not, are right-sided associations moderated by structural integrity of dominant left hemispheric language regions? Finally, can these results be explained mechanistically, extending the application of variable neuro-displacement?

METHODS

Participants

The same participants as Chapter 1 were included in this study. Twenty older, healthy native-English speakers were recruited (12 females, 46-76 years, mean age = 63.90, *SD* = 8.82). All participants had no history of neurological disease or language difficulties, normal/normal-to-corrected vision and no contraindications to MRI scanning. The Addenbrooke's Cognitive Examination Revised (ACE-R) was used for dementia screening, with a cut-off score of 88. The study was approved by a local National Health Service (NHS) ethics committee and all participants gave informed consent prior to participation.

Procedure

The training procedure was as reported in Gore et al., (2021; Chapter 1) with five stages: baseline naming assessments, three weeks of word training, post-training behavioural assessment, structural and functional imaging data collection and maintenance naming assessment. There were three sets of English language stimuli items; already known words, to-be-trained words and untrained words. All sets contained real-world items such as animals, food, tools and clothing. To-be-trained words and untrained words were rare, unfamiliar items. The known words could be named accurately with low word frequency. Stimuli sets were matched for psycholinguistic variables.

In baseline naming assessments, participants attempted to name all items. Items that could not be named were split between the to-be-trained and untrained word sets. Items that could be named were assigned to the known item set. Participants learned up to 50 native, novel words over three weeks from the to-be-trained item set during training. This training was both choice of cue training for item learning accuracy and a speeded

presentation task (Conroy et al. 2018) for item reaction time improvement. Both training methods are well known and commonly used speech and language therapies. Within 2 days of finishing training, participants were tested on all items without cues and scanned on the same day. Successfully named trained items were used during the fMRI naming task (trained vocabulary condition; M = 45 items). Participants were tested on all items five to six months post-scanning. No training took place between scanning and maintenance testing.

Neuroimaging acquisition and analysis

High-resolution structural T1-weighted MRI scans were acquired on a 3T Phillips Achieva scanner with a 32-channel SENSE head coil. Parameters were repetition time = 9ms, echo time = 3.93ms, flip angle = 8, FOV = 256mm, resolution matrix = 256 x 256, voxel size = 1.0 x 1.0 x 1.0mm, SENSE acceleration factor 2.5, total acquisition time = 575s with 150 contiguous slices. For functional data acquisition, a triple gradient echo EPI sequence was used to improve signal-to-noise ratio (SNR) in the anterior temporal lobes (ATLs), where traditionally there are issues of EPI signal dropout and distortion (Devlin et al. 2000; Poser et al. 2006; Halai et al. 2014). All functional scans were acquired using an upward tilt up to 45 degrees from the AC-PC line to reduce ghosting artefacts from the eyes into the temporal lobes. The sequence included 31 slices covering the whole brain with TR = 2.5s, TE = 12, 30 and 48ms, flip angle = 85 degrees, FOV = 240 x 240mm, resolution matrix = 80 x 80, and voxel size = 3.75 x 3.75 x 4mm. The full fMRI sequence is detailed in Chapter 1 (Gore et al., 2021).

Functional data were pre-processed and analysed as detailed in Chapter 1 (Gore et al., 2021). Structural MRI scans were pre-processed with Statistical Parametric Mapping software (SPM12; Wellcome Trust Centre for Neuroimaging, http://www.fil.ion.ucl.ac.uk/spm/). Firstly, the skull was stripped using an optimised brain extraction tool (OptiBET: Lutkenhoff, 2014). T1 structural images were normalised into Montreal Neurological Institute (MNI) space using a unified segmentation-normalisation procedure (DARTEL; Ashburner, 2007). Normalised whole brain images were smoothed with an 8mms full-width-half-maximum Gaussian smoothing kernel to account for global intra-subject differences. A group average grey matter explicit mask was applied.

The T1-weighted images were correlated with behavioural measures using voxel-based correlational methodology (VBCM, Tyler, Marslen-Wilson, & Stamatakis, 2005) using the PET/VBM module in SPM12 (Wellcome Trust Centre for Neuroimaging). This analysis correlated continuous signal intensity values in each voxel with the participants' behavioural measures. Several VBCM analyses were conducted: whole brain structural intensities versus naming trained item accuracy, reaction times (RT) and inverse efficiency scores (IES). All results were controlled for nuisance variables of age and total intracranial volume (TIV) estimates calculated by SPM12. The results were thresholded at $p = .005$ voxel level, $p < .05$ family-wise error (FWE) corrected cluster level.

Region of interest (ROI) analysis

A priori ROIs were constructed in the inferior frontal gyrus and IPL based on previous literature. The IFG is a critical area for speech production (Blank et al. 2002; Hickok and Poeppel 2007; Price 2010, 2012; Lazar and Mohr 2011). Grey matter intensity of the IPL is a neural marker for vocabulary acquisition success regardless of age (Green et al., 2007). The locations of the IFG (MNI: −46 28 10) and IPL (MNI: −47 −64 34) ROIs were taken from a Neurosynth meta-analysis of the term "language" as reported in Gore et al. (2021; Chapter 1). To explore the hypothesis that there may be associations between structure in the right hemisphere and behavioural performance, the left IFG and left IPL were mirrored to provide right IFG (MNI: 46 28 10) and right IPL (MNI: 47 −64 34) ROIs.

As T1 intensity ranges vary between participants due to scanner noise and skull thickness, T1 intensities were *z*-scored for each participant. For each voxel, T1 image means were taken from the voxel intensity, then divided by standard deviation (SD) for that individual's T1. The mean *z*-scored signal intensity for each ROI was regressed with behavioural and functional data. Reaction times (RTs) for ROI analyses were computed from the onset of stimuli and were *z*-scored to account for any variance due to time on task. Naming of known items was subtracted from naming of trained items to account for individual differences in general naming speed. ROI structural intensities were correlated with IES and mean ROI BOLD activity with Pearson correlation coefficients. This resulted in 12 contrasts adjusted for multiple comparisons by false discovery rate at an alpha level of .05 (Benjamini and Hochberg 1995). Two-stage

hierarchical multiple regressions were conducted in SPSS to test whether associations between right IFG structural intensities and IES/left IFG intensities were moderated by left IFG intensity. Right and left IFG intensities were mean centred, and the right and left IFG centred interaction term was generated.

RESULTS

Whole brain VBCM

VBCM analyses including grey and white matter identified the neural correlates of RT and IES (Figure 1, Table 1). There were no clusters that demonstrated a significant correlation between structural intensity and accuracy. There were two clusters where increased structural intensity correlated with reduced RT in the naming of newly trained items (controlling for general naming speed by subtracting average known word reaction times). A large cluster of increased structural intensity spanned across occipital cortices, right medial temporal areas and the right cuneus and inferior temporal gyrus.

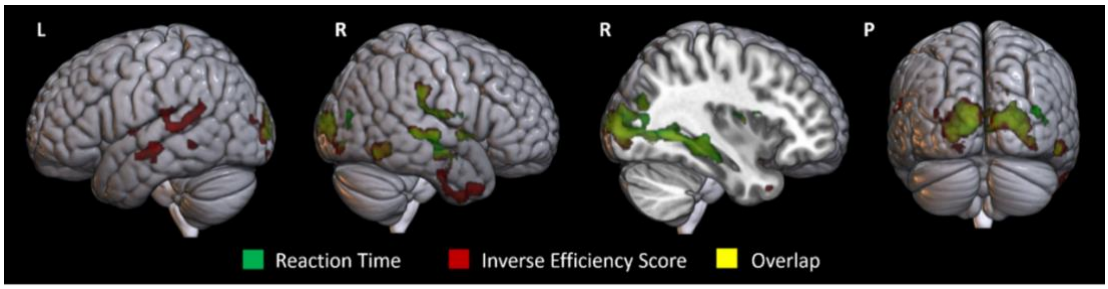


Figure 1. Whole brain structural correlates of naming performance for newly trained items. Areas of increased structural intensity correlated with quicker RTs (green) and better IES (red) with overlap (yellow). Age, total intracranial volume, and general naming speed were controlled for in both analyses. Images thresholded at $p <$ .005 voxel height, FWE cluster corrected $p < .05$. L; left, R; right, P; posterior.

The second cluster covered the Rolandic operculum, including the supramarginal and postcentral gyri spreading to the insula and superior/middle temporal gyri. VBCM analyses of IES showed correlations with structural intensity in the same areas as RT. Additionally, IES correlated with structural intensity in the right ATL and left posterior and mid temporal lobe.

Table 1. Areas of increased structural intensities correlated with behavioural performance measures.

VBCM Contrast	Cluster region	Peak region	Cluster size	Peak MNI x	y	z	*t*
RT	L/R occipital,	R vOT	3541	-10	-102	6	4.76
	R MTL	R Occipital		22	-70	-2	4.74
	R Rolandic op.,	R STG	1871	50	-30	22	6.16
	STG, insula	R SMG		58	-28	28	5.20
		R Rolandic op.		44	-16	18	4.39
IES	L/R occipital,	L cuneus	3425	-8	-104	6	5.50
	R ITG	R ITG		56	-52	-8	4.92
	R Rolandic op,	R STG	1064	48	-30	22	6.23
	R STG	R SMG		58	-28	30	5.15
	R temporal	R temporal pole	892	46	14	-40	5.34
		R MTG		64	-32	2	4.28
	L	L STG	872	-48	-34	24	8.47
	temporoparietal	L MTG		-62	-34	6	4.94
		L STG		-58	-46	14	4.84

Note: Structural intensity clusters significant at $p < .005$ voxel height and $p < .05$ FWE cluster correction. L; left, R; right, MNI; Montreal Neurological Institute, vOT; ventral occipitotemporal, MTL; medial temporal lobe, STG; superior temporal gyrus, ITG: inferior temporal gyrus; Rolandic op; Rolandic operculum, ATL; anterior temporal lobe, MTG; middle temporal gyrus.

Region of interest analyses

There was a significant negative correlation between right IFG structural intensities and inverse efficiency scores (IES; $r = -.553$, $p = .011$, adjusted $p = .013$, Figure 2b) of naming newly trained items (controlling for IES of already known items). In contrast, there was not a significant association between left IFG structure and trained > known IES ($r = -.118$, $p = .620$, adjusted $p > .999$). There was no significant correlation between right IPL structural intensities ($r = -.07$, $p = .769$, adjusted $p > .999$) or left IPL structure ($r = -.015$, $p = .966$, adjusted $p = .966$) and trained > known IES.

A two-stage hierarchical multiple regression was conducted to test whether the significant association between right IFG structural intensity and IES was moderated by left IFG structural intensity. A model including the right IFG and left IFG centred structural variables was significant, $R^2 = .318$, $F(2,17) = 3.960$, $p = .039$. The interaction term between right and left IFG intensities was added to the model and the three variables accounted for a significant proportion of the variance in IES, $R^2 = .498$, $F(3, 16) = 5.284$, $p = .010$. The interaction term significantly increased the amount of variance accounted for in the model, $\Delta R^2 = .180$, $\Delta F(1,16) = 5.729$, $p = .029$. Therefore, right IFG intensities were related to IES scores and left IFG intensities moderated that relationship. This interaction was visualised with predicted simple slopes to illustrate the conditional effects of right IFG intensities on IES at three levels of left IFG intensity (one SD below the mean, red; the mean, green; and one SD above the mean, blue; Figure 2*b*).

Simple slope analyses illustrate how the interaction term in the model relates to the actual data. There was variance between individuals in left IFG intensities. This data may be split into three groupings of IFG intensities; however this results in small sample size for each regression. Thus, to depict the moderation interaction, predicted simple slopes were generated (depicted on the right hand side of Figures 2b/c). Predicted simple slopes explore the conditional slope of a predictor when the moderator is held at a certain value (the mean, green; 1 SD above the mean, blue; and 1 SD below the mean, red). Therefore, predicted mean slope (green) may differ from the observed data slope (black), as each point on the observed data slope is an individual participant, with variation in left IFG intensities between participants. In the simple slopes analysis, each slope depicts a predicted individual with a moderator of a certain value (the mean, or 1 SD +/-). Confidence intervals indicate a prediction of how precisely each effect was estimated, which is influenced by factors such as sample size.

The relationship between the mean structural intensity and functional blood oxygenation level dependent (BOLD) activity in the newly trained > already known contrast of different ROIs was also explored. Right IFG structural intensities significantly correlated with right IPL activation ($r = -.642$, $p = .002$, adjusted $p = .012$) and left IPL activation ($r = -.066$, $p = .002$, adjusted $p = .024$) in the trained > known contrast. There were no further significant correlations between structural intensities and BOLD activity in the ROIs.

A further two-stage hierarchical multiple regression was conducted to test whether the significant association between right IFG intensity and right IPL activation was moderated by left IFG intensity. A model including the right IFG and left IFG centred variables was significant R^2 = .303, $F(2,17) = 3.695$, p = .046. The left-right structural IFG interaction term was generated and added to the model. The three variables also accounted for a significant proportion of the variance in IPL activation, R^2 = .473, $F(3, 16) = 3.526$, p = .039. The interaction term significantly added to the amount of variance accounted for by the model, ΔR^2 = .170, $\Delta F(1,16) = 5.161$, p = .0372. Therefore, left IFG intensities moderated the relationship between right IFG intensities and right IPL activation.

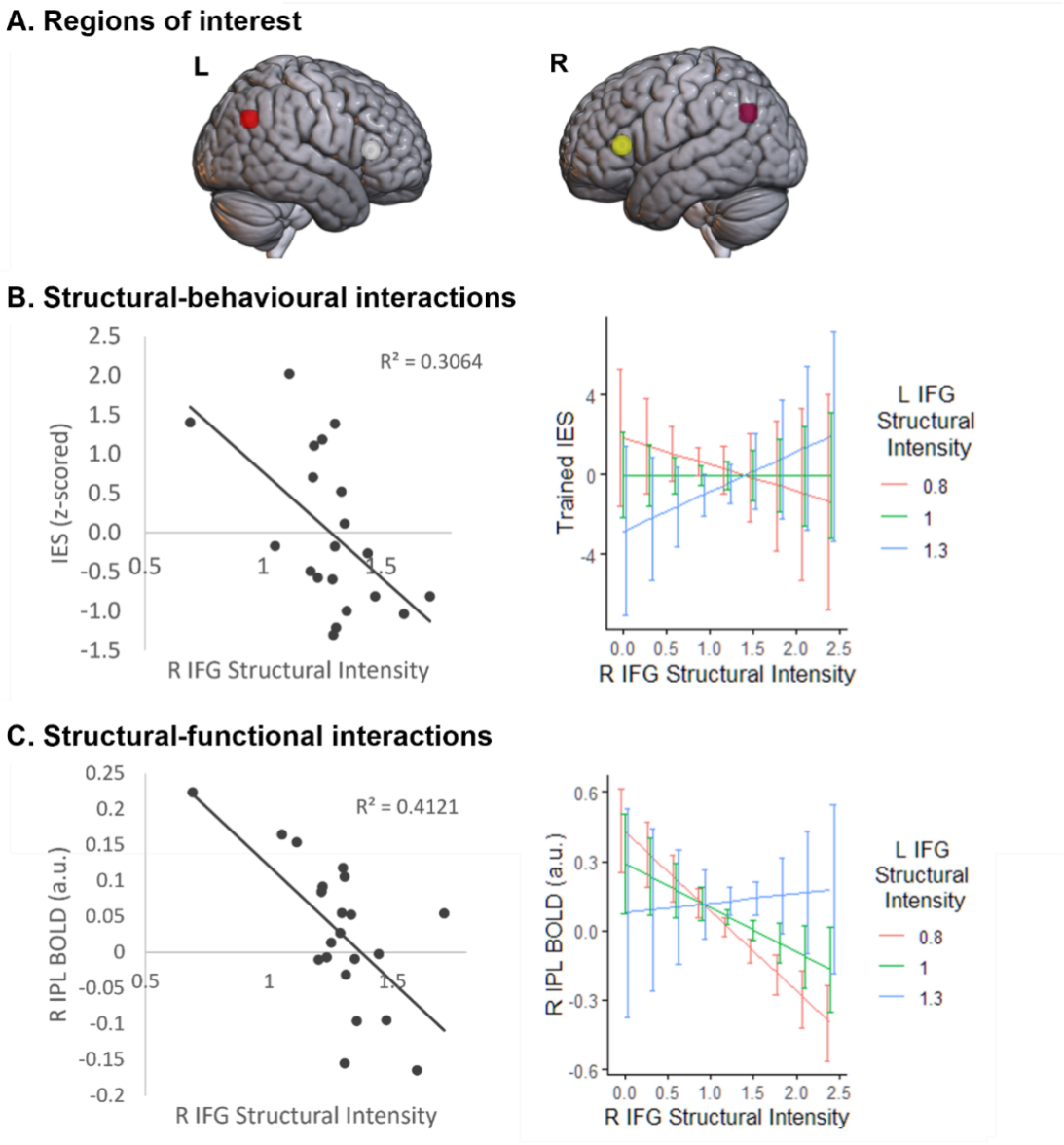


Figure 2. Associations between structural integrity and functional activity.

A. Locations of regions of interest; red; left inferior parietal lobe (IPL), white; left inferior frontal gyrus (IFG), yellow; right IFG, purple; right IPL. L; left, R; right. **B.** Right: Association between right inferior frontal gyrus (R IFG) structural intensities and newly trained item naming inverse efficiency scores (IES). **B.** Left: Predicted simple slopes of L IFG structural intensity moderation at three levels; one standard deviation (SD) below the mean (red), at the mean (green) and one SD above the mean (blue). **C.** Right: Association between R IFG structural intensities and R IPL BOLD in the trained > known contrast. **C.** Left: Predicted simple slopes of L IFG moderation at three levels.

DISCUSSION

This study explored whether the set of ideas termed variable neuro-displacement can be applied beyond aphasia to typical language learning. Specifically, we investigated the neurotypical framework of the structural correlates involved in word learning. The results demonstrated that increased structural intensity of right hemispheric language homologues was associated with better behavioural performance. Higher structural intensities in right IFG, a contralateral language homologue, correlated with better behavioural performance, and less activation of the right IPL when naming newly trained words. However, this association was moderated by the structural intensity of the left IFG. This suggests that greater structural intensity in left hemisphere language regions reduces the need to rely on their right hemisphere homologues when word learning is difficult and thus disrupts their relation to behaviour. The implications of this pattern of results are considered from a variable neuro-displacement perspective providing a mechanistic account of neurotypical language learning.

Areas that demonstrated a positive structural-behavioural relationship included areas of the network commonly found for speech production and picture naming (Price 2010, 2012). These included episodic (right hippocampus; Squire et al. 2004), semantic (right ATL; Lambon Ralph 2014) and visual regions, including the ventral occipitotemporal cortex. An area in the right middle SMG also displayed a positive association between structural intensities and quicker RTs and IES. This result aligns with previous literature, indicating IPL involvement in functional (Cornelissen et al. 2004a; Breitenstein et al. 2005; Pohl et al. 2017) and structural correlates of word learning (Lee et al. 2007). The anterior SMG is associated with phonological processing (Hickok and Poeppel 2007; Sliwinska et al. 2012), whereas the posterior IPL including the AG has been associated with semantic processing (Binder et al. 2009; Price 2012). Therefore, the SMG is well-positioned to connect phonological and semantic information about a word-picture referent during vocabulary acquisition.

The clusters of increased structural intensity associated with behaviour were largely right hemispheric, particularly correlates of RT when naming newly trained items. The incorporation of accuracy data in IES resulted in similar clusters of increased structural intensity but with the addition of some left hemispheric clusters. Accurate naming may require less capacity than naming an item both quickly and correctly, so accuracy may

relate solely to the dominant left hemispheric language network. In contrast, the more demanding task of quick response times requires the spare capacity within the right hemisphere, indicated by a right-sided network of intensity clusters. As a better IES score requires both high accuracy and short RTs it may best capture different aspects of difficulty and thus, reflects both right and left hemisphere language regions.

Multiple possible inferences can be made when relating increased structural intensities and behaviour. A large body of research suggests skill acquisition and experience can alter brain structure (Draganski et al. 2004; Lövdén et al. 2013; Wenger et al. 2017). Therefore, the results may indicate that more proficient word learners had greater grey matter densities in domain-specific regions, due to experience and practise. However, it is important to consider that the participants in this study were older to provide an age-matched comparison for aphasia studies. There are well documented age-related changes in neural structure, resulting in tissue loss of grey (Good et al. 2001; Resnick et al. 2003; Raz et al. 2005) and white matter (Sowell et al. 2003; Giorgio et al. 2010). Increased intensities in this population may reflect a lack of atrophy compared to other individuals. Age was covaried to account for these effects. However, it is still an important consideration as biological age may not accurately reflect all variance in the change through time (i.e., some people may see the effects of ageing quicker or more slowly due to genetic or lifestyle factors). There are documented increasing levels of bilaterality in ageing language systems as compared to a typically left lateralised language network in younger adults (Tyler et al. 2010; Hoffman and Morcom 2018).

These results raise questions as to why the language production system is architecturally bilateral, yet many processes are dominated by the left hemisphere. One potential explanation for these findings is that bilateral systems provide resilience to the effects of damage. In computational models, unilateral damage reduces relevant activity in the damaged hemisphere, which allows for the intact hemisphere to dominate and maintain performance (Schapiro et al. 2013). However, this does not explain why a single hemisphere supports many language processes. It is possible that the language network is overall bilateral, whereby both hemispheres can support function. But, to save metabolic energy, the level of activation is titrated against performance demand, which can be affected by difficulty of task and damage to the network as described in the set of ideas termed variable neuro-displacement (Stefaniak et al. 2020).

As such, structural intensities of the left IFG moderated these structural-behavioural and structural-functional relationships. Only when the left IFG was sufficiently lacking in structure was the right IFG an important factor. The left IFG is a critical region in the dominant language hemisphere involved in speech production (Blank et al. 2002; Price 2010, 2011). This result indicates that bilaterality in typically left dominant processes is dependent on neural reserve and structural integrity of the most efficient parts of the network in neurotypical participants. Only when the critical left IFG was deficient in either capacity or efficiency through increased performance demands and/or age-related structural changes was the structure of the right contralateral homologue important. This moderation of associations is consistent with the notion that right hemispheric structures become more important in aphasics with larger left hemispheric lesions (Heiss et al. 1999; Saur et al. 2006; van Oers et al. 2010; Szaflarski et al. 2013).

Conclusions

In conclusion, the results of this study support the utility of a potential extension of the idea of variable neuro-displacement beyond applications in the aphasic literature to normal language learning. Increased right hemispheric structural intensities correlated with better learning of novel words and decreased parietal lobe activation, however, these associations were moderated by structural intensities in the dominant left hemispheric language network. A mechanistic explanation for these results may be provided by the principles of variable neuro-displacement with increased engagement of right hemisphere areas in response to increased task demands and/or a system that lacks capacity or efficiency. Questions remain as to whether aphasic patients would manifest right hemispheric activation during re-learning of words and novel word learning. These questions are considered in Chapters 4 and 5.

CHAPTER 4: WORD RE-LEARNING IN POST-STROKE APHASIA ENGAGES NEUROTYPICAL NEURAL MECHANISMS UP TO A CRITICAL DAMAGE POINT

ABSTRACT

Post-stroke aphasia can cause catastrophic deficits in communication and severely reduce quality of life. Speech and language therapy (SLT) is an effective tool in improving post-stroke outcome. SLT may induce typical recovery mechanisms in stroke, with associated functional reorganisation. However, SLT may also drive neurotypical learning mechanisms. In healthy individuals, word learning fits within the Complementary Learning Systems framework. In this "reverse translation" study, 16 chronic post-stroke aphasia patients undertook well-studied SLT interventions for three weeks. Patients completed a subsequent fMRI picture naming task, which probed (i) treated items, (ii) already known items and (iii) untreated/unknown items. Overall, the results demonstrated that patients relied on similar underlying neural mechanisms as healthy controls when re-learning items, fitting within the CLS framework. Increased levels of remaining left inferior frontal gyrus (IFG) activity was associated with quicker responses, and the level of damage to this critical language region, moderated the relationship between hippocampal activity and the naming of treated items. In patients with relatively preserved left IFG, hippocampal activity was associated with lower accuracy and slower responses. However, patients with over 50-60% damage to the left IFG displayed the opposite pattern, with more hippocampal activity when naming associated with quicker responses. Hippocampally-dependent patients may be either (i) in the initial learning phase due to a lack of retained partial representations or (ii) unable to shift the division of labour to the damaged cortical language systems.

INTRODUCTION

A common consequence of cerebral stroke is language difficulties or 'aphasia'. Communication is vital to quality of life (Code, 2003). Thus, successful treatment of aphasia is paramount. Speech and language therapies (SLT) have been intensively tested and there is substantial evidence that therapy can be effective (for a review cf. Nickels, 2002). However, the underlying neural correlates of successful and unsuccessful therapy remain unclear. There are benefits to understanding successful and unsuccessful therapy, as the latter may help with improving therapies with better targeting and stratification of patients.

Previous literature has explored treatment-induced changes in brain activity, with some consistency across broad theories of standard recovery processes. SLT may boost these standard processes and induce neuroplasticity (Berthier and Pulvermüller 2011). In some studies, recovery is dependent on the reconstitution of language systems in undamaged tissue around the lesion, resulting in left perilesional activity (Hillis et al. 2006; Saur et al. 2006; Meinzer and Breitenstein 2008; Fridriksson 2010; van Oers et al. 2010; Abel et al. 2015). Other studies report a shift of language function to the contralateral hemisphere, with increased activity in right language homologues (Blasi et al. 2002; Leff et al. 2002; Turkeltaub et al. 2012). However, there is debate as to which patterns of reorganisation are advantageous to language outcome in stroke. In speech production, the contribution of right hemispheric key language homologues has been considered both beneficial (Blasi et al. 2002; Leff et al. 2002; Crinion and Price 2005) and detrimental (Naeser et al. 2004; Winhuisen et al. 2005) in recovery.

Alternatively, therapy may promote some of these standard recovery mechanisms and also encourage others. There is previous evidence of training-induced functional changes in domain-general brain networks, particularly networks underpinning short-term memory. In healthy individuals, neural processes associated with word training fits within the Complementary Learning Systems theory (CLS; McClelland et al., 1995, 2020; Davis et al., 2009; Kumaran et al., 2016). The CLS model proposes that the hippocampus and medial temporal lobes (MTL) support initial rapid, sparse knowledge acquisition. Slower, more in depth representations are supported by the neocortex, allowing for generalisation across episodes. Over time, knowledge consolidation is driven by a gradual shift in the division of labour between these complementary

systems. In Gore et al., (2021; Chapter 1) healthy participants were trained on native, novel items. The results of this study fit within the CLS theory model, with newer less consolidated items associated with activation of hippocampus and episodic regions, and fully consolidated already known items associated with critical neocortical language region activity. Does aphasia treatment encourage typical knowledge acquisition mechanisms during the re-learning of items?

In this reverse translational study, we aimed to elucidate the underlying neural correlates of successful speech and language therapy. Critically, does word re-learning in post-stroke aphasia fit within the framework of word learning in healthy older adults? Or does a certain level of critical damage cause migrations from this typical learning framework? To explore these questions, we gave patients three weeks of speech and language therapy on items they could not name in prior naming assessments. Post-therapy, patients performed a picture naming task in the functional MRI scanner naming: treated items, untreated items they could name prior to scanning, and untreated items they could not name prior to scanning. The design of this study followed the procedure of a previous study in healthy controls (Chapter 1; Gore *et al.*, 2021). Thus, we have some understanding of what neurotypical novel naming looks like and how this compares to naming of well consolidated, established items. This background knowledge allowed us to compare when and if patients' underlying neural activity reflects that of healthy controls.

Patients were asked to name/re-learn vocabulary that they had previously been able to retrieve. We predicted there were three possible outcomes. (i) If enough of the patient's typical language systems were intact, then the naming of these items should mirror item naming in healthy controls. (ii) If representations of the item name exist but were weakened, re-learning could strengthen the representations and processing in the typical (retained) language system. This process may be supported by 'new learning' via the episodic system, similar to when controls learn new vocabulary (Gore et al. 2021). (iii) If language systems were more compromised, then the remaining representations may be more severely damaged and thus, (a) compensatory language-related systems may be engaged, or (b) if the representations were so weakened that re-learning is closer to novel vocabulary learning, then episodic systems would be dominant, as they are in healthy individuals during the initial phase of new vocabulary learning.

METHODS

Participants

Sixteen participants were recruited from the North West of England (7 female, age range = 39-73 *M* = 57.5) with acquired language production impairment following a single left hemispheric stroke (either ischaemic or haemorrhagic) at least one year prior to the study (see Figure 1 for lesion overlap and Table 1 for demographic and selected neuropsychological data). Participant's aphasia was classified using the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983). All were native English speakers, with normal or normal-to-corrected vision and hearing, with no other significant neurological conditions or contraindications to MRI scanning. Participants gave informed consent prior to participation. Ethics were approved by a local National Health Service (NHS) ethics committee and conducted per the ethical principles as detailed in the Declaration of Helsinki. Structural MRI data from an age and education matched healthy control group (n = 22, 10 female, M = 68.21) were used for lesion identification (Seghier et al. 2008).

Background neuropsychological assessments

All participants were tested on an extensive neuropsychological/aphasiological battery assessing language and cognitive abilities (Butler, Lambon Ralph, & Woollams, 2014). In addition to the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983) this battery consisted of (i) naming tests, including the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983) and the 64-item Cambridge naming test; (ii) syntax-level tests including the spoken sentence comprehension task from the Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2005); (iii) semantic tasks from the 64-item Cambridge Semantic Battery (Bozeat et al. 2000) and the 96-trial synonym judgement test (Jefferies et al. 2009); (iv) speech production tests from the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992); (v) cognitive tests including the forward and backward digit span (Wechsler 1955), Raven's Coloured Progressive Matrices (Raven 1962) and the Brixton Spatial Rule Anticipation Task (Burgess and Shallice 1997). Patient's demographic details are displayed in Table 1, alongside results of the BNT to demonstrate that there was a word-finding deficit. An overlap map of participants' left hemispheric lesions is shown in Figure 1.

Table 1. Patient's demographic information ordered by lesion volume.

Patient ID	Age (years)	Gender	Education (years)	BDAE classification	Time post-stroke (years)	Lesion volume (voxels)	BNT
P1	57	F	16	Anomia	6	175	61.67
P2	73	M	11	Anomia	6	3136	76.67
P3	63	M	19	Anomia	4	4234	50.00
P4	53	F	11	Anomia	5	6630	78.33
P5	59	M	16	Anomia	5	6974	55.00
P6	49	M	11	Anomia	6	8163	43.33
P7	52	F	19	Anomia	6	8810	88.33
P8	66	F	12	Anomia	9	9383	75
P9	57	F	11	Anomia	3	10027	71.67
P10	62	M	17	Broca	10	11179	76.67
P11	58	F	11	Anomia	5	12767	50.00
P12	54	F	11	Anomia	6	13080	78.33
P13	55	M	11	Anomia	8	14681	85
P14	55	M	11	Broca	3	17465	35
P15	50	M	11	Broca	9	20681	41.67
P16	58	M	13	Anomia	6	22298	38.33

Note: BDAE = Boston Diagnostic Aphasia Examination; BNT = Boston Naming Test.

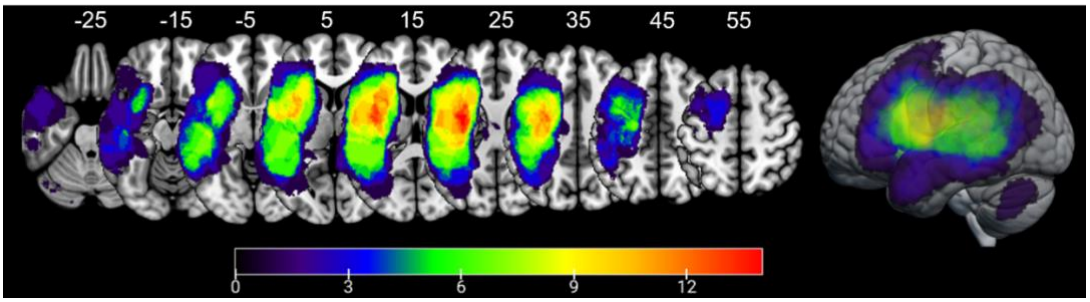


Figure 1. Lesion overlap map of the 16 participants on axial slices. Colour bar indicates the number of participants included in the overlap. Numbering above slices indicates z MNI coordinate. In neurological convention, left is left.

Pre-treatment naming assessments and stimuli

The design mirrored that used in Gore et al. (2021) with healthy controls. Patients were tested on all stimuli items three times prior to training. There were three sets of items: already known, to-be-trained and untrained. 200 items were taken from the International Picture Naming Project and matched with discernible, coloured images with a white background to build the already known and to-be-trained item sets. Items were selected with high word frequency, short reaction times (<1000ms) and high accuracy (85-100%). Items that were accurately named on all occasions were added to that individual's already-known set. Items that could not be named on one or more occasions were randomly added to either to-be-trained or untrained sets. Final sets were only 50 items in total to allow for psycholinguistic matching between sets. Participants who could accurately name over 80% of all items prior to training were provided with a more challenging set of to-be-trained items (3 participants in total). This set contained unfamiliar, rare items with low word frequency names, drawn from the British National Corpus (BNC Consortium, 2007). Images from the already known, treated and untreated sets were phase scrambled for the baseline task.

Anomia treatment

Patients received a laptop to practice naming the to-be-treated items for 45 minutes a day, four days a week, for three weeks. Patients completed, on average, a total of 7.3 hours of naming treatment. During the first two weeks, participants received cue training. In the third week, participants received speeded training.

The cue training item number was incrementally increased, beginning with 10 items. When 70% accuracy was reached on these 10 items, another 10 items were added incrementally up to the total 50 items. Participants first viewed each item picture and item name in orthographic and audio form. The patients repeated the name out loud of all items in the current set. The cue training was designed with standard error-reducing speech and language therapy (Abel et al. 2005; Conroy, Sage, and Lambon Ralph 2009), but with a choice of cue to provide a self-determined approach. An item picture was shown, and patients had an option of cues or to proceed with naming. For example, for a picture of a butterfly, the cue options were a written semantic description "A flying insect with large colourful wings", an initial phonemic cue 'bu', initial and second phonemic cue 'butter' or finally, the whole word cue 'butterfly'. All cues were delivered orthographically and audibly. After each naming attempt, the whole correct word was given. Patients then answered a semantic question "Is it living?".

The third week of training was given in repeated increasingly-speeded presentation (RISP; Conroy, Sotiropoulou Drosopoulou, Humphreys, Halai, & Lambon Ralph, 2018). Patients were instructed that they needed to name the item before the computer. When participants reached a speeded naming success rate of 70%, the timing was incrementally decreased from 3s to 1.4s, to 1s. When participants beat the 1s target for 70% of items, the set size was increased by 10 items and the time-to-beat returned to 1.8s.

Neuroimaging acquisition

High-resolution structural T1-weighted MRI scans were acquired on a 3T Philips Achieva scanner using an eight-element SENSE head coil. The parameters were as follows: repetition time = 9.0ms, echo time = 3.93ms, acquired voxel size = 1.0 x 1.0 x 1.0 mm squared, matrix size = 256 x 256, slice thickness = 1mm, flip angle = 8 degrees, FOV = 256mm, inversion time = 1150ms, with 150 contiguous slices and a SENSE acceleration factor of 2.5.

A triple gradient echo EPI sequence was used for functional imaging to improve signal-to-noise ratio in the anterior temporal lobes (ATL; Kundu et al., 2012, 2017). In single echo imaging, there is signal dropout and distortion in the ATL (Poser et al. 2006; Halai et al. 2014, 2015b). Functional scans were also acquired with a 45-degree tilt from AC-PC to reduce ghosting artefacts in the temporal lobes. The functional sequences

consisted of 31 slices covering the whole brain, TR = 2.5s, TE = 12, 30 and 48ms, flip angle = 85 degrees, resolution matrix = 80 x 80, FOV = 240 x 240mm and voxel size = 3.75 x 3.75 x 4mm.

Stimuli were presented during scanning using E-Prime 2.0, with block order pseudo-randomised and optimised for statistic power using OptSeq (http://surfer.nmr.mgh.harvard.edu/optseq/). Spoken in-scanner responses were recorded using a fibre optic, noise-cancelling microphone for fMRI (FOMRI; Optoacoustics). Patients practised speaking with as little movement as possible prior to scanning to reduce motion artefacts.

There were two tasks during imaging acquisition, one of which is the focus of a separate study. For this study, a blocked design picture naming task with four conditions was used; already known, treated, untreated and baseline. In the known condition, patients named pictures they could consistently name prior to treatment (e.g., pencil). In the treated condition, patients named newly treated items (e.g., penguin). If patients could not remember an item name, or the item was novel or phase scrambled they responded: "Don't know". The second task, the focus of a separate study, involved patients making semantic judgements in three blocked conditions: treated, untreated and baseline. In the treated and untreated conditions, patients answered the questions "Is it living?" in response to orthographic item names. In the baseline task, participants responded "Up" to an ascending alphabetical sequence "ABCD" or "Down" to a descending alphabetical sequence "DCBA".

In both tasks, each trial lasted 3700ms. A fixation cross was presented for 700ms, followed by the image to be named for 3000ms. Each 11.1s block consisted of three trials. Eight jittered length rest blocks were included per run, with an average length of 11.1s. There were two 7 minute and 4-second-long blocks per specific task.

Neuroimaging preprocessing and analysis

T1-weighted structural images were pre-processed in FSL, version 6.0.0 (Woolrich et al. 2009), using the fsl_anat pipeline (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/fsl_anat), excluding FAST tissue segmentation. A supervised lesion segmentation algorithm (LINDA; Pustina et al., 2016) was used for lesion identification. T1-weighted images were normalised to Montreal Neuroimage Institute (MNI) space and LINDA transforms were applied. The functional data were despiked and slice time corrected per echo in the

AFNI neuroimaging suite (v19.2.10; 3dDespike; 3dTshift; Cox, 1996; Cox and Hyde, 1997). Motion correction parameters were calculated from the first slice of the first echo, then this transformation was applied to all three echoes. A subject-specific, LINDA-generated T1 mask was transformed to functional space and applied to each echo. Multi-echo data were optimally combined using the T2* combination method (Posse et al. 1999) in tedana (Kundu et al., 2012; The tedana Community et al., 2021). Optimally combined data were normalised to MNI space using LINDA transforms. Smoothing was performed on this data using FSL, using 8mm FWHM equivalent sigma. Statistical whole brain analyses were performed non-parametrically using SnPM13 (http://warwick.ac.uk/snpm). Implicit masking defaults were set to '-Inf' to remove the arbitrary analysis threshold and avoid lesion voxels being masked out for all participants. Region of interest analyses were performed using SPM12 (Penny et al. 2007) and MarsBaR (Brett et al. 2002). Two-step multiple regression and simple slope analyses were performed in R version 4.0.5 (RStudioTeam 2020) using the emmeans package.

Regions of interest were defined based upon previous literature. Five *a priori* regions of interest were defined to include key semantic and episodic network areas. The left inferior frontal gyrus (IFG; MNI: -46 28 10), a key area of damage in post-stroke aphasia is considered critical in speech production (Blank, 2002; Hickok and Poeppel, 2007; Price, 2012). The right IFG homologue (rIFG; MNI: 46 28 10) was included to investigate right-sided homologue compensation theories (Rosen et al. 2000b; Naeser et al. 2005; Heiss and Thiel 2006). A left ventral anterior temporal lobe (vATL; MNI: -36 -15 -30) was taken from a key semantic cognition reference (Binney et al. 2010). For the episodic memory network, a left hippocampal ROI (MNI: -28 -14 -15) was defined based on previous literature of initial hippocampal activation during word learning (Breitenstein et al. 2005; Davis et al. 2009). In addition, an ROI in the left inferior parietal lobe was included (IPL; MNI: -47 -64 34) due to consistent activation of this region in episodic processing (Cabeza et al. 2008; Humphreys et al. 2020). Pearson's *r* correlations were used to explore associations between mean ROI BOLD activity and behavioural performance measures (accuracy and RT). Critical tests of significant differences between correlations were adjusted for multiple comparisons using the Benjamini-Hochberg procedure, FDR = 0.05 (Benjamini and Hochberg 1995).

To explore the interaction between critical language region damage and performance/functional activity, a further left IFG ROI was created. This ROI included left IFG pars opercularis and left IFG pars triangularis, as designated by the Automated Anatomical Labelling (AAL) atlas (Rolls et al. 2020).

Behavioural performance analysis

Individual behavioural performance was probed for differences between the three pre-training test scores using a McNemar's test. Pre-treatment scores were averaged per participant as baseline performance was stable across the group and individuals. Repeated-measures *t*-tests were performed for each stimuli type, between the timepoints of pre- and post-treatment for accuracy and RT. Two repeated-measures analyses of variance (ANOVA) were conducted, one for accuracy performance and one for RT. These ANOVA analyses included within-subject main factors of Time (pre-treatment, post-treatment) and Treatment (treated, untreated and already known). Two-tailed *p* values and an alpha level of 0.05 were applied for all statistical tests.

RESULTS

Behavioural performance

There was no significant difference in the proportion of accurate responses between the three baseline pre-tests across the group ($p = .88$). As pre-treatment performance showed no significant change across timepoints, baseline pre-tests were averaged, to provide a single representative sum of pre-treatment baseline performance. There was no significant difference in accuracy ($t = 1.67$, $p = .12$) or RT ($t = -1.76$, $p = .10$) for untreated items between timepoints (Figure 2a). Patients had significantly more accuracy ($t = 5.18$, $p < .0001$) and faster RTs ($t = -7.63$, $p < .0001$) when naming treated items after treatment (Figure 2b). Patients were also significantly more accurate naming known items ($t = 3.31$, $p = .004$) but did not have a significant difference in naming speed ($t = -2.08$, $p = .06$; Figure 2c).

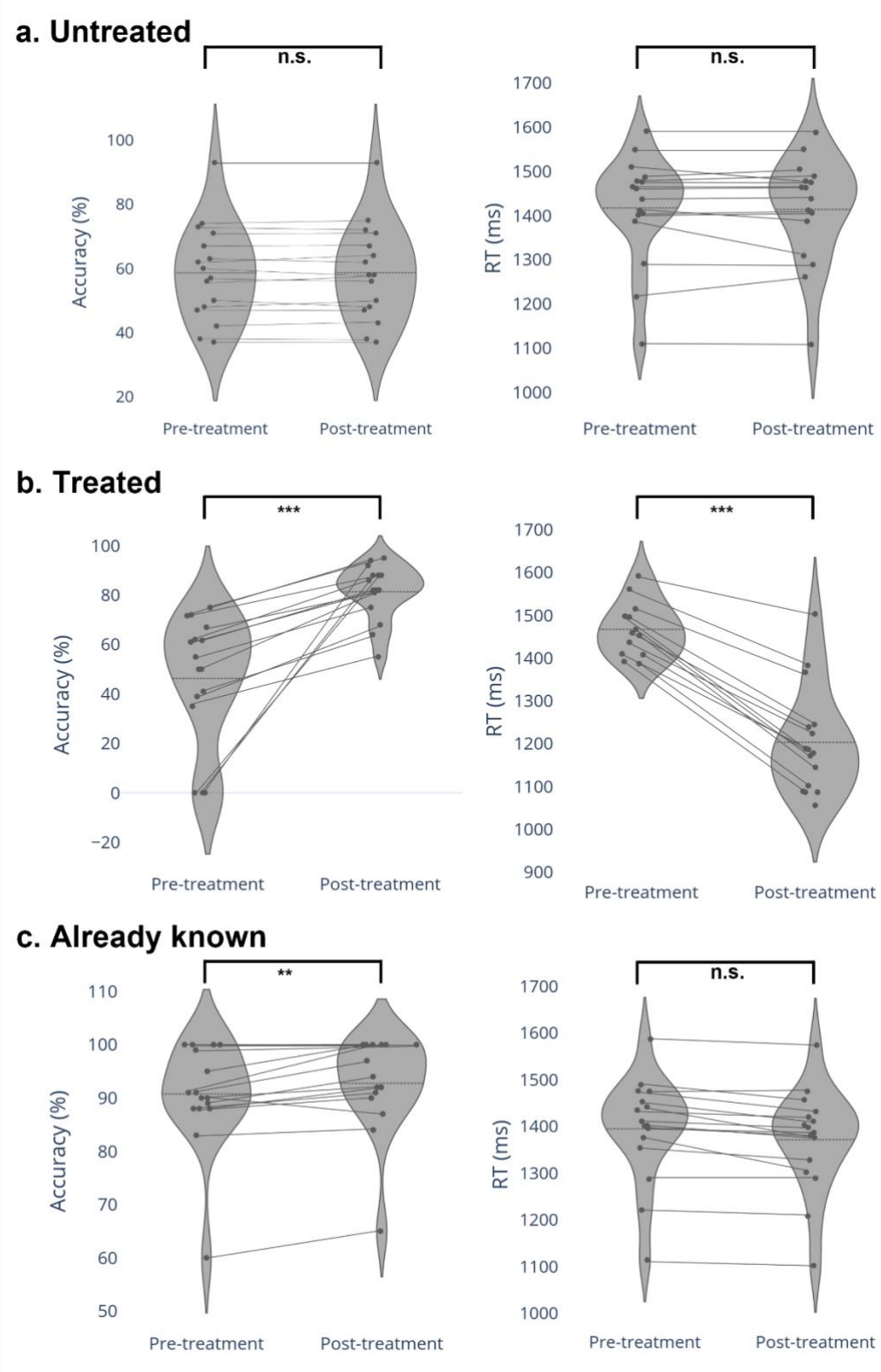


Figure 2. Accuracy and reaction times (RT) of naming a) untreated items, b) treated items and c) already known items. Dotted line represents the mean. Solid lines between points represent individual patient differences in time points. Group differences significant at $p < .0001$ and $p < .005$ are denoted with '***' and '**' respectively.

Accuracy

A repeated-measures 2 x 3 ANOVA revealed a significant main effect of Time (F (1,15) = 90.35, $p < .001$), with greater accuracy post-treatment. There was also a significant effect of Treatment (F (2,30) = 118.31, $p < .0001$) with greater accuracy when naming newly treated than untrained or known items. There was a significant Time x Treatment interaction (F (2,30) = 35.48 $p < .0001$). Naming of treated items was more accurate post-treatment (81%) than pre-treatment (46%). This unstandardised effect size (35%) was significantly greater than the mean difference (1%) between naming of untreated items pre-treatment (58%) and post-treatment (59%; $p < .0001$). The unstandardised effect size was also significantly greater than the mean difference (2%) between naming of already known items pre-treatment (90%) and post-treatment (92%; $p < .0001$).

RT

There were also significant main effects of both Time (quicker RTs post-treatment; F (1,15) = 24.084, $p < .001$) and Treatment (faster RTs when naming newly treated than untrained or known items; F (2,30) = 27.307, $p < .0001$). There was also a significant Time x Treatment interaction for RT (F (2, 30) = 17.507, $p < .001$). Naming of newly treated items was faster (1203ms) than pre-treatment naming (1467ms). This unstandardised effect size (264ms) was significantly greater than the mean difference (4ms) between untreated items pre-treatment (1417ms) and post-treatment (1413ms; $p = .002$). Additionally, the unstandardised effect size was also significantly greater than the mean difference (23ms) between already known items pre-treatment (1394ms) and post-treatment (1371ms; $p = .002$).

Whole brain fMRI contrasts

The results of whole brain analyses are reported in Table 2. Multiple contrasts were performed; treated > already known, treated > unknown, already known > unknown and the inverse of these contrasts. There were significant clusters for treated > unknown and already known > unknown (Figure 3). In the treated > unknown contrast, there was a significant cluster of activation over the right superior temporal gyrus, right rolandic operculum and the right postcentral gyrus. There was also a further cluster in the dorsal occipital cortex spanning bilateral lingual gyri and the cerebellum. In the already known > unknown contrast, there was a similar pattern of activation. The right-sided cluster spanned the right anterior parietal and superior temporal cortex areas. The posterior

cluster centred on the bilateral lingual gyri, including posterior left parahippocampal cortex.

Table 2. Significant clusters of activity during naming of treated, already known and untreated items.

Contrast	Region of activation	Peak region	Cluster size	Peak MNI			Pseudo *t*
				x	y	z	
Treated > untreated	R STG, postcentral gyrus	R STG	3947	65	-14	11	5.26
		R postcentral		54	-6	13	5.05
		R STG		64	-6	4	3.86
	Occipital cortex, cerebellum	R cerebellum	24708	17	-59	-23	5.15
		L calcarine		2	-86	-5	4.84
		L cerebellum		-13	-62	-13	4.33
Known > untreated	R SMG, R STG, R Heschl gyrus	R SMG	2855	61	-35	29	4.84
		R STG		65	-16	12	4.35
		R STG		64	-7	7	3.76
	Bilateral lingual gyrus, left MTL	L fusiform	9039	-31	-39	-11	4.71
		R lingual		5	-79	-7	4.65
		L calcarine		2	-85	0	4.39

Note: Clusters significant at $p < .001$ voxel height and $p < .05$ FWE cluster correction. Up to 3 strongest peaks listed per cluster. Pseudo-t statistics are computed with smoothed variance in SnPM. R, right; STG, superior temporal gyrus; SMG, supramarginal gyrus; MTL, medial temporal lobe.

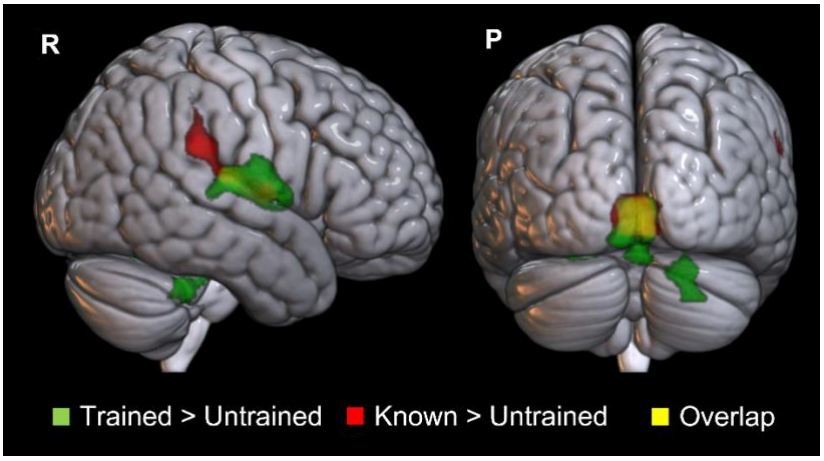


Figure 3. Whole brain clusters of increased activation during picture naming. Picture naming of trained > untrained (green) and known > untrained (red) contrasts, with overlap (yellow). Image thresholded at voxel level $p < .001$, cluster corrected FWE $p < .05$. R; right, P; posterior.

Region of interest analyses

To explore the core hypothesis that neural processes underlying word re-learning in aphasia would be similar to those that support word learning in healthy controls (e.g., CLS theory) six a priori regions of interest (ROIs) were defined. These included language-semantic network areas (left inferior frontal gyrus (IFG), right IFG and left anterior temporal lobe; ATL) and episodic network areas (bilateral hippocampi and left inferior parietal lobe). Percentage damage to 8mm spherical ROIs is listed in Table 3. An initial exploratory analysis was performed correlating BOLD activity of these ROIs with naming reaction times (RT) and accuracy. There were no significant correlations between behaviour and ROI activity in the already known > untreated contrast.

Table 3. Percentage damage to the regions of interest (ROI) per participant. All ROIs refer to an 8mm spherical region, except AAL L IFG which refers to the whole left IFG pars opercularis and pars triangularis.

	Patient ID															
ROI	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**
L IFG	84	-	-	-	100	6.8	-	-	-	61	82	-	22	-	-	55
R IFG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
L ATL	-	-	12	-	-	-	-	-	-	-	-	-	-	-	-	-
L Hipp	-	42	13	-	-	-	-	2	-	-	-	-	6	-	-	1
L IPL	87	-	-	-	-	-	79	-	-	-	-	62	34	-	-	-
AAL L IFG	95	89	10	37	94	38	28	89	-	88	99	97	95	-	4	47

Note: "-" indicates no lesion overlap with the ROI. L, left; R, right; IFG, left inferior frontal gyrus; ATL, anterior temporal lobe; Hipp, hippocampus; IPL, inferior parietal lobe; AAL, Automated Anatomical Labelling.

In Gore et al. (2021; Chapter 1), the underlying neural correlates of word learning fit within the CLS model. To test the hypothesis that patients would follow this typical healthy learning framework, we explored whether treated > unknown (whereby patients named pictures of treated items versus responding to unknown and untreated items) behavioural correlations differed significantly from already known > unknown items. In healthy participants, left hippocampal activity was associated with worse performance, whereas left IFG and left ATL activity was associated with better performance (Chapter 1). For the patients in this study, there was a strong negative correlation between left hippocampal activity and accuracy in the treated > unknown condition (r = -.517, p = .039, Figure 4b), which was significantly different to the weak *positive* correlation of left hippocampal activity and accuracy in the already known > untreated condition (r = .363, p = .166) using Fisher's r-to-z transformation (z = -2.43, p = .008, adjusted p = .024). There was a positive correlation between left IFG activity and naming accuracy in the treated > unknown condition (r = .540, p = .043, Figure 4*b*), which was significantly different to the weak negative correlation of left IFG activity and naming accuracy in the already known > untreated condition (r = -.025, p = .929) using r-to-z transformation (z = 2.19, p = .014, adjusted p = .021). There were no further significant correlations between accuracy and ROI BOLD activity, including in the right IFG r = .041, p = .881.

Associations between region of interest BOLD activity and naming RTs were also explored. There was a strong positive correlation between left hippocampal activity and longer RTs in the treated > unknown condition (r = .576, p = .020. Figure 4*c*), which was significantly different to the weak positive correlation in the already known > unknown condition (r = -0.011, p = .971; z = -1.65, p = 0.04, adjusted p = .04). There were no further significant correlations between RT and ROI BOLD activity, including in the right IFG (r = .229, p = .260).

To test a further hypothesis, that neural changes may deviate from normal word learning paradigms dependent on levels of damage to critical regions, these brain-behaviour associations were further probed. The strongest difference in brain-behaviour correlations between treated and already known items was between left hippocampal BOLD activity and RT. Hierarchical multiple regression was performed to probe this association. All covariates were mean centred, and the

interaction term was generated. As there was little variation in the level of damage to the 8mm left IFG ROI, percentage damage to the whole left IFG was considered. In the first step, variables of left hippocampal activity and percentage left IFG damage were entered as predictors, accounting for a non-significant amount of variance in RTs (R^2 = .340, F (2, 13) = 3.345, p = .067). In the second stage, the interaction term between left hippocampal activity and left IFG damage was entered accounting for a significant proportion of variance (R^2 = .765, F (1, 12) = 5.655, p = .012, b = -.508, t = -2.669, p = .02).

Predicted simple slopes were generated to illustrate how this interaction term relates to the actual data, at three levels of left IFG damage, one standard deviation (SD) above the mean (blue; 98%), one SD below the mean (pink; 18%) and the mean (yellow; 58%; Figure 2*c*). Predicted simple slopes explore the conditional slope of a predictor when the moderator is held at a certain value. When an individual has a low level of IFG damage, greater hippocampal activity was associated with longer RTs. In contrast, when an individual has a higher level of left IFG damage, greater hippocampal activity was associated with shorter RTs.

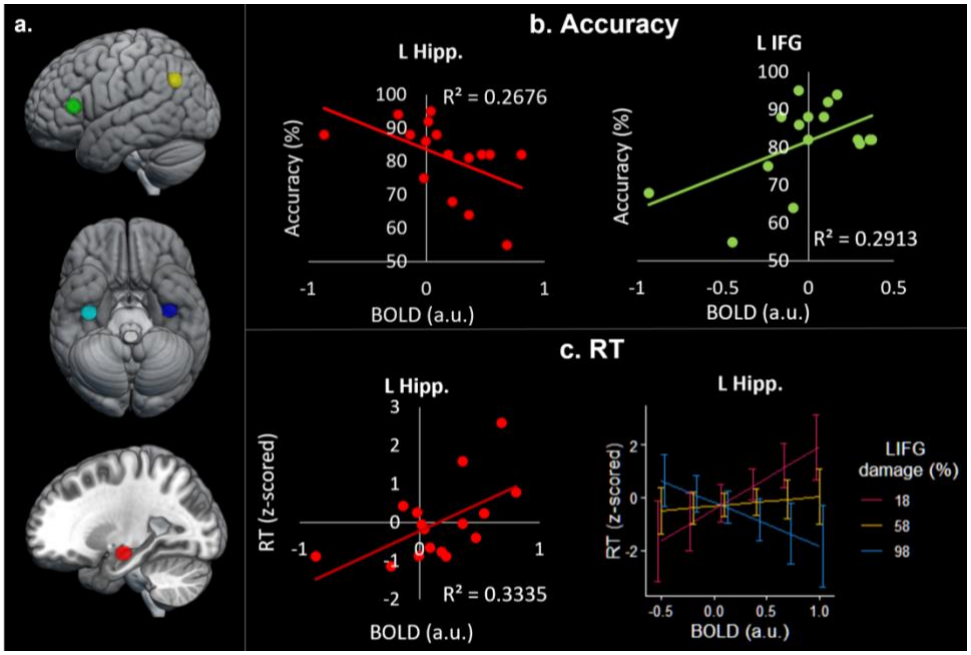


Figure 4. Correlations between regional activation and behavioural performance. (A) Regions of interest: green = left inferior frontal gyrus (IFG); yellow = left inferior parietal lobe; cyan = left anterior temporal lobe (ATL); blue = right ATL; red = left hippocampus (Hipp.). (B) Correlations between treated > untreated BOLD activity and accuracy. (C) Left: correlation between treated > untreated left hippocampal BOLD activity and RT. Right: Predicted simple slopes of the association between treated > untreated left hippocampal BOLD activity and RTs at three levels of left IFG damage (pink = 1 SD below mean, yellow = mean, blue = 1 SD above mean).

DISCUSSION

We investigated whether underlying neural mechanisms of word re-learning in post-stroke aphasia fit within a framework of healthy older word learning. Previous studies of treatment-induced changes in neural activity have demonstrated that therapy may i) drive standard recovery processes and ii) drive domain-general processes. Healthy participants have previously undertaken the same task as the patients in this study, providing a healthy word learning framework (Gore et al. 2021a). In healthy controls, word learning follows the CLS theory of general knowledge acquisition (Davis and Gaskell 2009). The episodic left hippocampus supports initial learning (and associated worse performance). In contrast, consolidated knowledge (and better performance) is associated with typical speech

production areas such as the left IFG and ATL (Gore et al., 2021; Chapter 1). In this study, patients with post-stroke aphasia undertook the same procedure. Patients were trained on 50 items for 3 weeks before undertaking a naming task during fMRI. In region of interest analyses, patients' brain-behaviour correlations followed a similar pattern as seen in healthy word learning. Increased hippocampal activity during naming of treated words was associated with worse performance (i.e., slower RTs and lower accuracy). In contrast, increased activity in a critical area of the language network, the left IFG, during naming of treated words was associated with quicker RTs. Crucially, the association with hippocampal activity was moderated by degree of damage to the left IFG. Overall, increased episodic network activity was associated with worse performance and increased language-semantic network activity with better performance. These brain-behaviour associations mirror the pattern found in healthy individuals and fit within the CLS model. Below, we discuss these results in more detail and consider how they might fit with theories of recovery and therapy.

The whole brain results demonstrated an overlapping cluster of increased activity in both the already known > untreated and treated > untreated contrasts, centred on the right posterior superior temporal gyrus (pSTG). Over half of the patients had lesion overlap within the left pSTG (9; Figure 1). The right pSTG is consistently activated in controls and patients with aphasia during speech production tasks (Stefaniak et al. 2021). Indeed, there was right pSTG activation for controls performing the same task (Chapter 1; Gore *et al.*, 2021). Thus, naming in post-stroke aphasia may be facilitated by typical language regions. However, speech production in controls is associated with bilateral activation of pSTG and areas of the left prefrontal cortex (Price 2010; Gore et al. 2021; Stefaniak et al. 2021), indicating a reduced level of left hemispheric activation in patients for naming.

Episodic systems engagement was modulated by the degree of damage to the left IFG. The division of labour between memory systems of the CLS model might vary with aphasia severity because if mildly aphasic there are more retained, partial representations of the to-be-learned items left in the damaged cortical systems. Thus, therapy drives the refinement of those partially damaged representations (i.e., the balance of labour already favours cortical over hippocampal contributions). However, if patients are much more severely aphasic (with more cortical damage), it

is more similar to learning the items from scratch. As in the CLS model, these patients' initial learning processes are hippocampally-dependent.

Alternatively, if patients with low levels of left IFG damage are less aphasic (with better baseline naming performance) and thus more similar to controls, patients can consolidate the items well in cortical areas and reduce hippocampal contributions. Those with more IFG damage may be more aphasic and have fewer cortical language systems to pick up the longer-term consolidation of re-learned items. Therefore, therapy performance remains reliant upon the hippocampus. This would suggest that, in a CLS framework, individuals can only shift divisions of labour to cortical systems if those cortical systems are intact enough to support longer-term learning. If the cortical language systems are too damaged, re-learning depends on hippocampal memory systems. If these theories were to hold, then it may be possible to substitute the moderator of left IFG damage with either i) other critical language regions or ii) baseline naming ability and demonstrate the same interaction term.

These results imply that it may be important to consider the level of damage to critical language regions during stratification of patients to therapy approach. Patients with more damage to critical language regions and therefore, more hippocampal dependence, may show worse learning retention over time. To test this hypothesis, further research is needed with post-treatment testing to determine maintenance of re-learned items. However, if hippocampally-dependent participants are either i) in an initial phase of learning due to a lack of retained partial representations or ii) unable to shift the division of labour to the damaged cortical language systems, a longer-term treatment strategy may be crucial to treatment outcomes. Overall, larger sample sizes are needed to obtain a fuller picture of the interaction between language and episodic networks in post-stroke aphasia with increased variation in lesion profile and language deficit.

We observed small gains in the already known and untreated item sets. Although a central aim of therapy intervention is generalisation of treatment effects to untreated words, most interventions result in change primarily on treated items (Nickels 2002; Best et al. 2013). There are important considerations regarding the source of these gains in this study. Patients were tested on all items a total of five times. This re-assessment alone may be sufficient to increase performance (Howard et al. 2015).

However, previous literature suggests that generalisation to untreated items via cueing therapy may be specific to a minority of patients with relatively intact semantic processing and deficits only in the phonological encoding stages of speech production (Best et al. 2013). The patients in this study were generally mildly aphasic. Except for one patient, all had only minor semantic deficits. This may account for the increases in performance of the already-known words in this study.

Conclusions

The results of this study demonstrate that word re-learning post-stroke follows the typical healthy word learning framework, up until approximately 58% damage to the critical left IFG region. The opposite effect was demonstrated after critical damage to the left IFG, with episodic activity associated with better performance. Hippocampally-dependent patients may be either (i) in the initial learning phase due to a lack of retained partial representations or (ii) unable to shift the division of labour to the damaged cortical language systems.

CHAPTER 5: NEURAL UNDERPINNINGS OF NOVEL, NATIVE WORD LEARNING IN POST-STROKE APHASIA

ABSTRACT

This study investigated the neural underpinnings of native, novel word learning in post-stroke aphasia. Four participants with post-stroke aphasia were trained on up to 50 previously unknown native words over three weeks. Participants named already known, trained and untrained/unknown items during a picture naming fMRI task. The neural correlates of word learning in post-stroke aphasia mirrored that of previously seen in 20 healthy controls, which fit well within the Complementary Learning Systems model that proposes initial hippocampal episodic encoding followed by cortical consolidation. Similar to controls, naming of newly learned items was associated with increased activation of both episodic and language regions. However, unlike controls, naming of already known items was also associated with episodic and language region activation in PWA with more extensive lesions. Additionally, PWA maintained a variable amount of learned item names within the group, comparatively similar to controls. Participants with reduced structural connectivity (within lesioned tissue and extending beyond into intact perilesional white matter) in the arcuate fasciculus and inferior fronto-occipital fasciculus retained fewer items over a period of three months. The results of this case series suggest that although the functional correlates of word learning were largely normal, structural abnormalities in the dorsal and ventral pathways was associated with the longevity of learning retention, with implications for the success of therapeutic intervention.

INTRODUCTION

Approximately one third of stroke survivors suffer from language production and/or comprehension deficits, known as “aphasia” (Berthier 2005; Engelter et al. 2006). Patients with aphasia (PWA) post-stroke form a heterogeneous group with a broad range of varied, contrastive language impairments. Aphasia is traditionally defined by categorical subtyping with differing neuropsychological and lesion profiles (Broca 1861; Wernicke 1874). However, word finding difficulties, or “anomia”, are a common problem across PWA and impact everyday communication (Goodglass and Wingfield 1997). Thus, a significant portion of aphasia treatment research is devoted to naming therapy, which can provide effective rehabilitation, although success is variable across individuals (Nickels 2002). An understanding of the neural mechanisms supporting successful therapy is key to improving rehabilitation outcome prediction, tailoring therapy to the individual to maximise gains.

The first step to optimising how language is re-learned in the damaged brain is to understand which neural mechanisms support word learning in the healthy brain. With a baseline of healthy learning, comparisons can be made as to how novel words are learned in a damaged brain. An influential theory of knowledge acquisition is the Complementary Learning Systems (CLS) theory (McClelland et al. 1995, 2020; Kumaran, Hassabis, and McClelland 2016). This theory posits that the neocortex and hippocampus have complementary properties which allow for learning. Specifically, the hippocampus is involved in fast learning of sparse non-overlapping representations of episodes. This process is considered complementary to slow learning of highly overlapping representations in the cortex, allowing for generalisation across episodes. Davis and Gaskell (2009) proposed that word learning fits within the CLS model. There is growing evidence for this proposal from computational modelling (McClelland 2013; Kumaran, Hassabis, and Mcclelland 2016; Schapiro et al. 2017; McClelland et al. 2020), neuroimaging studies (Cornelissen et al. 2004; Breitenstein et al. 2005; Davis et al. 2009; Takashima et al. 2014; Gore et al. 2021) and neuropsychological data (Scoville and Milner 1957).

Most recently, we (Gore et al., 2021; Chapter 1) gave healthy older adults the same training regime on previously unfamiliar native words as those used in the present study. These participants therefore act as matched controls for the chronic stroke

aphasic patients considered here. The results aligned with the CLS model of memory. Newly learned items were supported by a combination of regions associated with episodic memory (including left hippocampus) and language-semantic areas that support established vocabulary (left inferior frontal gyrus and left anterior temporal lobe). There was a shifting division of labour across these two networks in line with items' consolidation status; faster naming was associated with more activation of language-semantic regions and lesser activation of episodic memory regions. This framework can be used to compare the underlying neural processes of novel word learning in healthy participants and a series of PWA.

Few studies provide PWA with novel, native words to learn, as therapy is generally focussed on reinstatement of previously familiar words. There is already some, albeit limited, evidence that PWA can indeed learn novel pseudowords (Gupta et al. 2006) and novel native words (Tuomiranta et al. 2012, 2015; Tuomiranta, Grönroos, et al. 2014; Dignam, Copland, et al. 2016) with varying degrees of success. The small number of studies of novel words in PWA may be due to higher frequency words being considered a better target in terms of their everyday utility. However, novel word learning is a useful tool as, milder patients with anomia can find repeated intervention on simple high-frequency items to be a relatively boring task leading to drop-out (Brady et al. 2016). Secondly, novel items do not have pre-existing semantic representations. Existing word-referent associations for familiar items may influence the already opaque re-learning process, given that anomia can be driven by phonological and/or semantic deficits (Schwartz et al. 2006).

Additionally, existing knowledge varies significantly between participants, which can cause difficulties in establishing stable pre-treatment baselines (Howard et al. 2015). Finally, novel word learning allows for comparison with healthy controls. Speech and language therapy of familiar items is an unsuitable task for individuals without word finding deficits, whereas novel word learning is suitable for both healthy controls and PWA. This means that the learning of novel native words in age matched controls by Gore et al. (2021) provides an effective comparison between the normal language learning system and the damaged system when contrasted with the learning of those with aphasia.

Within the small number of studies of novel word learning in PWA, few include associated neuroimaging data beyond lesion profile. However, Coran et al. (2020) performed a novel word learning study in three PWA with associated diffusion-weighted imaging. This study linked partial sparing of the arcuate fasciculus (AF) with better word learning performance. Due to the AF connecting Broca and Wernicke's area (Catani et al. 2005), and evidence of AF damage resulting in conduction aphasia (Anderson et al. 1999), the AF has been considered a key tract in language processing. Further, there is evidence that strength of structural and functional AF connectivity is crucial for word learning in healthy participants (López-Barroso et al. 2013). To our knowledge, only one study has used explored novel word learning in PWA with functional imaging. Tuomiranta et al. (2015) performed an online learning study, whereby a patient with global aphasia viewed items from the Ancient Farming Equipment paradigm (Laine and Salmelin 2010) paired with associated novel names. The patient in this study successfully learned 10 novel items with associated increases of left hippocampal activation. These results indicate that word learning in PWA may align with the CLS model. However, further research with larger samples and long-term word learning is needed to draw conclusions.

There are methodological issues to consider when using single-subject fMRI, such as statistical power limitations. However, recent work has indicated that speech-related brain activity may be reliable in single-subject analyses (Frankford et al. 2021). Furthermore, there are associated benefits with using single-subject fMRI. Individuals with post-stroke aphasia form a heterogenous group, with graded differences in lesion profiles, post-injury deficits and pre-morbid cognitive ability (Ingram et al. 2020) and neural structures (Forkel et al. 2014). Thus, underlying neural mechanisms of word learning may also vary in PWA. Case series allow for interpretation at the subject level across a series of related patients to identify similarities and differences modulated by specific variables (Schwartz and Dell 2012). As prediction of therapeutic outcomes must be at the level of the individual, there is an advantage to case-series analysis alongside larger group studies (Howard 2003).

The aim of this case-series study was to investigate the neural correlates of novel word learning in post-stroke aphasia. Four PWA learned the names for up to 50

novel English words and associated semantic information. Analyses were performed at the individual level to identify patterns across variables. The specific aims of this study were as follows: i) To replicate previous evidence of novel word learning in post-stroke aphasia; ii) To determine whether neural underpinnings of novel word learning in aphasia fit within the CLS memory framework; iii) To explore whether structural integrity of key language white matter tracts, particularly the AF, is associated with word learning in aphasia.

METHODS

Participants

Four stroke survivors (3 female) with chronic post-stroke aphasia were recruited from a database of participants held by the Neuroscience and Aphasia Research Unit (NARU) at the University of Manchester. These participants were recruited from communication support groups across the North West and referrals from health practitioners. The selection criteria for patients were correct naming of over 80% of 200 items drawn from the International Picture Naming Project (IPNP; Chapter 4). The four participants were classified as having anomic aphasia on the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983), hence their aphasia was relatively mild. All were right-handed, native English speakers who suffered a single left hemispheric stroke resulting in language difficulties at least one year prior to the study. They had no contraindications to MRI scanning, severe apraxia of speech or confirmed or suspect diagnoses of dementia. Informed consent was obtained prior to participation in the study and ethics were approved by a local National Health Service ethics committee, in accordance with the Declaration of Helsinki. High resolution T1 images were used from an education- and age-matched healthy control group (Chapter 1) for lesion identification (Pustina *et al.*, 2016; Figure 1).

Case descriptions:

P1

P1 was a 48-year-old female shop manager at the time of the study, with 13 years of formal education. P1 presented with a large (18948 voxels) left hemispheric middle cerebral artery (MCA) infarct lesion, encompassing the entire left inferior frontal gyri and insula, spanning across areas of the left motor cortex and anterior inferior parietal lobe (Figure 1a). At the time of participation, P1 was 7-years post-stroke. P1 had moderate non-fluent speech production and reported struggles with function words and longer words but presented with good comprehension skills. P1 also had mild right-sided hemiparesis affecting mobility.

P2

P2 was a 57-year-old female at the time of the study, with 12 years of formal education. P2 suffered a large (9383 voxels) left hemispheric (MCA) infarction which involved the posterior inferior frontal gyrus, inferior motor cortex and anterior portions of the parietal cortex (Figure 1b). P2 suffered complete loss of speech production and reading abilities post-stroke, alongside right-sided hemiplegia. At the time of participation, 14-years post-stroke, P2 was well-recovered with fluent speech, only mild word finding difficulties and mild right-hand weakness. P2 had also recently learned to read again, as a consequence of considerable effort to regain this skill which had been important pre-morbidly.

P3

P3 was a 58-year-old musician at the time of the study, with 16 years of formal education. P3 presented with a moderate (6974 voxels) left hemispheric lesion (Figure 1d), 5-years-post MCA ischemic stroke. The lesion encompassed the anterior parietal lobes including the left inferior parietal lobe, angular gyrus and supramarginal gyrus. In addition, the lesion included the left lateral temporal cortex including the entire middle and superior temporal gyri and areas of dorsolateral inferior temporal gyrus. P3 was generally well recovered but had deficits of sentence level processing with mild conversational agrammatism and word-finding difficulties.

P4

P4 was a 47-year-old female at the time of the study, with 16 years of formal education and had worked as an orthopaedic nurse. At the time of participation, P4 was 7-years-post-stroke. P4's stroke led to a small lesion (175 voxels) of the left supplementary motor area (Figure 1c). In addition, automated lesion identification demonstrated abnormal areas of the cerebellum. Post-stroke, P4 suffered right-sided hemiplegia and severe aphasia. Since, P4 recovered with generally mild word finding difficulties with fluent speech, writing and comprehension. However, P4 still suffered from poor speech production in group settings.

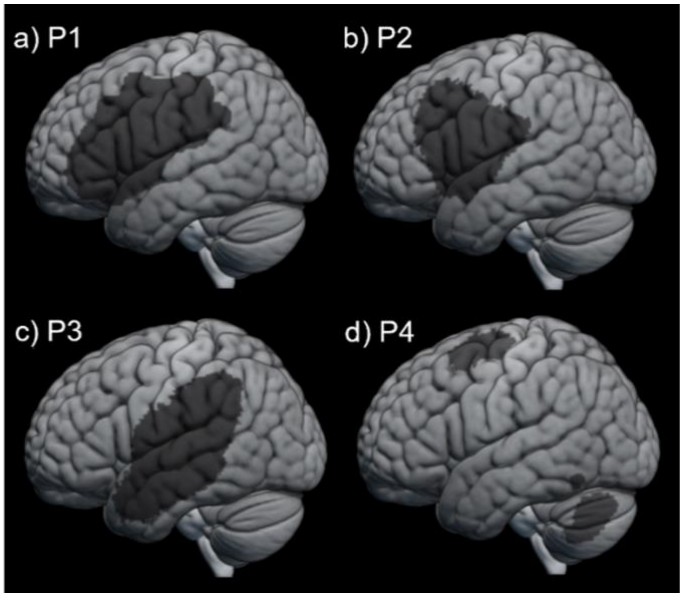


Figure 1. Left-hemispheric LINDA-generated normalised lesion binary maps. Lesion depicted by dark grey, overlayed on a standard brain for each participant.

Behavioural assessment battery

PWA were tested on an extensive neuropsychological battery assessing both predominantly language but also some other cognitive abilities (Table 1; Butler, Lambon Ralph, & Woollams, 2014). In addition to the BDAE (Goodglass & Kaplan, 1983) the neuropsychological battery consisted of (i) naming tests including the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983) and the 64-item Cambridge naming test; (ii) syntax level tests including the spoken sentence

comprehension task from the Comprehensive Aphasia Test (CAT; Swinburn, Porter, & Howard, 2005); (iii) semantic tasks from the 64-item Cambridge Semantic Battery (Bozeat et al. 2000) and the 96-trial synonym judgement test (Jefferies et al. 2009); (iv) speech production tests from the Psycholinguisitic Assessments of Language Processing in Aphasia (PALPA; Kay, Lesser, & Coltheart, 1992); (v) cognitive tests including the forward and backward digit span (Wechsler 1955), Raven's Coloured Progressive Matrices (Raven 1962) and the Brixton Spatial Rule Anticipation Task (Burgess and Shallice 1997).

Table 1. Demographic information and background language assessments.

Task	**P1**	**P2**	**P3**	**P4**
Gender	F	F	M	F
Age	48	66	59	57
Formal education (years)	13	12	16	16
Time post (years)	7	9	5	6
BDAE classification	Anomia	Anomia	Anomia	Anomia
BNT (%)	84.38	88.33	78.33	88.33
Cambridge Naming (%)	60.00	93.75	87.50	93.75
Immediate word repetition (%)	92.50	100	91.25	100
Delayed word repetition (%)	88.75	98.75	98.75	100
Word-Picture-Matching (spoken) (%)	100	100	95.31	98.44
Camel and Cactus (%)	90.63	79.69	92.19	84.38
Raven's Matrices (%)	91.67	91.67	83.33	94.44
Lesion volume (voxels)	18948	9383	6974	175

Note: BDAE = Boston Diagnostic Aphasia Examination; BNT = Boston Naming Test.

Procedure

PWA undertook the same procedure as healthy participants reported in Gore et al. (2021; Chapter 1). There were five stages to the study. Baseline naming assessments were given on three occasions before three weeks of novel word training. Within 2 days of training completion, PWA were tested on all items and underwent a picture naming task during functional imaging. Three months after scanning, without any ongoing training, PWA were tested on learned items to assess maintenance.

Pre-treatment naming assessments and stimuli

There were four sets of stimuli: already-known, trained, untrained and a baseline set. Stimuli sets included real-world items such as tools, mammals, fish, birds, clothing, and foods. The known item set contained familiar items drawn from the International Picture Naming Project, e.g. dragonfly, xylophone, hairdyer. The untrained and trained item sets contained unfamiliar items with very low word frequency names, e.g. echidna, dilruba, binnacle. These items were drawn from the 100-million-word text, British National Corpus (BNC Consortium 2007). Stimuli were counterbalanced across trained and untrained sets. The baseline set contained phase scrambled images for the known, trained and untrained item sets. PWA were tested on all stimuli items at 3 timepoints prior to training with at least a day interval (Howard et al. 2015). Items named correctly from the known set on all three occasions were added to the participant's already-known set. Items which could not be named on any occasion were added to that participant's untrained/to-be-trained sets.

Novel word training

PWA learned novel words and related semantic information at home for 45 minutes a day, four days a week, for three weeks. During the first two weeks PWA received computed-based cue training, a commonly used speech and language therapy (Nickels 2002; Abel et al. 2005). Items were trained in incremental sets and started with 10 items. These novel picture items were displayed with the associated name presented orthographically and audibly. Participants repeated these items overtly. Then, each item was presented with a choice of cue including increasing phonological cues and a semantic cue. Participants could pick any number of cues before naming the item. After naming each item, participants were asked "Is it

living?" to encourage semantic learning. Once 70% accuracy was achieved naming the 10 items, another 10 items were added incrementally up to the full 50 novel items. In the third week, participants received repeated increasingly speeded therapy (Conroy et al. 2018). Participants attempted to name a presented learned item before the computer did so, to increase naming speed. When participants reached a speeded naming success rate of 70%, the timing was incrementally decreased from 3s to 1.4s, to 1s. When participants beat the 1s target for 70% of items, the set size was increased by 10 items and the target time to beat returned to 1.8s.

Participants completed on average a total of 10.5 hours (*SD* = 1.78) of naming practice. Post-training naming assessments were performed in the absence of cues. Successfully named items were used during the fMRI naming task (trained vocabulary condition; *M* = 42 items, *SD* = 1.73). The fMRI session took place on the same day as the post-training assessment. Reaction times were calculated from the onset of stimulus.

Neuroimaging acquisition

All imaging was performed on a 3 T Philips Achieva scanner (Philips Medical Systems, Best, Netherlands) using an 8 element SENSE (Sensitivity Encoding Spin Echo) head coil. High-resolution, whole brain structural images were acquired with a weighted inversion recovery procedure with 3D acquisition in 150 contiguous slices, repetition time (TR) = 8.4ms, echo time (TE) = 3.8ms, acquired voxel size = 1.0 x 1.0 x 1.0 mm squared, 256x256 matrix, slice thickness = 1mm, flip angle = 8 degrees, SENSE acceleration factor = 2, field of view (FOV) = 256mm, inversion time = 1150ms, total scan acquisition 574s.

Functional data were acquired with a triple gradient echo EPI sequence was used to improve signal-to-noise ratio in the semantic anterior temporal lobes (ATL), an area often subject to signal dropout (Poser et al. 2006; Halai et al. 2014, 2015a; Kundu et al. 2017). In addition, functional scans were acquired with a 45-degree tilt from AC-PC to reduce ghosting artefacts from the eyes in the temporal lobes. The functional sequences consisted of 31 slices covering the whole brain, TR = 2.5s, TE = 12, 30 and 48ms, flip angle = 85 degrees, resolution matrix = 80 x 80, FOV = 240 x 240mm and voxel size = 3.75 x 3.75 x 4mm. Stimuli were presented via E-Prime 2.0. Block order was pseudo-randomised using OptSeq

(http://surfer.nmr.mgh.harvard.edu/optseq/). In-scanner overt verbal responses were recorded using a fibre optic, noise-cancelling microphone for fMRI (FOMRI; Optoacoustics).

Diffusion-weighted data were acquired using a pulsed gradient spin echo (PGSE) echo planar imaging (EPI) sequence implemented with TE = 54ms, Gmax = 62 mT/m, half scan factor = 0.679, 112 x 112 image matrix reconstructed to 128 x 128 using zero padding, reconstructed resolution 1.875 mm x 1.875 mm, slice thickness 2.1 mm, 60 contiguous slices, 43 non-collinear diffusion sensitisation directions at b = 1200 s/mm^{-2} (Δ = 29.8 ms, δ = 13.1 ms), 1 at b = 0, SENSE acceleration factor 2.5. TR was dependent on individual participant heart rate due to cardiac gating. 44 temporally spaced volumes with different direction diffusion gradients were acquired. Each diffusion gradient orientation had phase encoding in the right-left and left-right directions and thus reversed phase and frequency encode direction (Embleton et al., 2010). Acquisitions were cardiac gated using an index finger peripheral pulse unit. Cardiac gating resulted in different length TR and scan duration; however, the approximate acquisition time was 28 minutes.

During functional imaging for this study participants performed a picture naming task, naming four sets of items: already-known, newly trained, untrained and baseline images. Participants responded overtly to all items. When participants could not recall an item name, or, when responded to untrained/baseline images, participants were instructed to respond “don’t know.” Each trial lasted 3700ms and consisted of a fixation cross for 700ms, followed by the target image for 3000ms. Three items were presented per block and each block lasted 11.1s. 8 jittered rest blocks were included with an average length of 11.1s. With 32 task blocks and 8 rest blocks, the total run time was 7 minutes and 4 seconds. There were 2 functional runs performed.

Neuroimaging preprocessing and analysis

T1-weighted structural images were pre-processed in FSL, version 6.0.0 (Woolrich et al. 2009), using the fsl_anat pipeline (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/fsl_anat) with default settings, except excluding FAST tissue segmentation in the initial stage. Lesion profiles were identified using LINDA (Pustina *et al.*, 2016), a supervised

lesion segmentation algorithm. Structural T1-weighted images were normalised to Montreal Neuroimage Institute (MNI) space and LINDA transforms were applied.

In the AFNI neuroimaging suite (v19.2.10; Cox, 1996; Cox and Hyde, 1997) functional data were despiked and slice time corrected per echo (3dDespike; 3dTshift). Motion correction parameters were calculated from the first slice of the first echo and these motion transformations were applied to all three echoes. A subject-specific, LINDA-generated T1 mask, including lesioned voxels was transformed to functional space and applied to each echo. The three echo-times were optimally combined using the T2* combination method (Posse et al. 1999) in tedana (Kundu *et al.*, 2012; The tedana Community *et al.*, 2021). Optimally combined triple-echo data were normalised to MNI space using LINDA transforms and smoothed in FSL, using 8mm FWHM equivalent sigma. Statistical whole brain analyses were performed using SPM12 (http://warwick.ac.uk/snpm) in a first and second level General Linear Model. Implicit masking defaults were set to '-Inf' to remove the arbitrary analysis threshold and avoid lesioned voxels being masked out for all participants. Significant threshold was set to $p < .001$, cluster level $p < .05$ (FWE-corrected for multiple comparisons across the whole brain. Anatomical labelling was determined using the Automated Anatomical Labeling atlas (Rolls et al. 2020).

Diffusion-weighted data were pre-processed using FSL's diffusion pipeline. Brain tissue was extracted from the b = 0 data (BET; Smith 2002). Data was collected with reversed phase-encode blips, resulting in pairs of images with distortions going in opposite directions. From these pairs the susceptibility-induced off-resonance field was estimated (Andersson et al. 2003) and the two images were combined into a single corrected one (topup command). Estimation and correction of eddy current induced distortions and motion artefacts were corrected using eddy_openmp (Andersson and Sotiropoulos 2016). A diffusion tensor model was fitted at each voxel with the dtifit command to obtain mean diffusivity (MD) and fractional anisotrophy (FA) values. A per participant inclusive mask was created using the MD map and thresholding 50% of the range of non-zero voxels. This removed voxels in cerebrospinal fluid and lesioned space, where sensible fibres or tracking is not expected. Masks were visually inspected. A probabilistic diffusion model using bedpostx (Jbabdi et al. 2012) for crossing fibres was applied to inclusive voxels.

Default parameters were used except for the following: fibres per voxel = 3, model = 3 (zeppelin axially symmetric tensor), burn-in period = 3000. To map long-range connectivity differences between PWA and the healthy control group, anatomical connectivity maps (ACMs; Bozzali et al. 2011) were computed for each PWA and control using FSL's probtrackx2 function with 50 streamlines per voxel. The output reflects how connected each voxel is to the rest of the brain. ACMs were normalised into MNI space with a transformation matrix between the B0 in diffusion space and the structural T1 image. These normalised images were smoothed with an 8mm full-width half-maximum Gaussian kernel to account for intersubject variability. Each participant's ACM was entered into a two-sample t-test in SPM12, against the 20 control ACMs.

RESULTS

Behavioural performance

Participant's behavioural performance data is listed in Table 1. PWA successfully learned novel, native words with an average 84% accuracy, successfully learning an average of 42 items with little variability (SD = 2.828). Healthy control participants learned the same novel words with 90% accuracy, learning an average of 45 items. Total of learned items was not significantly different between PWA and healthy controls (t = -0.78, p = .442). Participants maintained an average of 61.5% of items three months post-training, which was not significantly different to controls (t = -1.06, p = .29; Table 2).

Table 2. Behavioural performance of naming already known and newly trained items.

	P1		P2		P3		P4	
Task	**Pre**	**Post**	**Pre**	**Pre**	**Post**	**Post**	**Pre**	**Post**
Known accuracy	90	91	92	94	96	96	100	100
Known RT	1412	1401	1321	1208	1534	1328	1224	1102
Trained accuracy (%)	0	82	0	82	0	84	0	88
Trained RT (ms)	-	1328	-	1221	-	1287	-	1114
Trained maintenance (%)	-	20	-	5	-	71	-	100

Note: Pre = pre-training, Post = post-training, RT = reaction time, ms = milliseconds. Known accuracy and reaction time (RT) detail an average over three pre-training baseline assessments. Post-training refers to the post-training behavioural assessment immediately prior to scanning. Maintenance indicates percentage of successfully trained items retained after three months.

Functional neuroimaging

Whole brain single-subject functional analyses for the contrasts trained > unknown and known > unknown are displayed in Figure 2 and listed in Tables 3 and 4. There were no significant clusters of activity in a group analysis.

P1

In the trained > unknown contrast, a large cluster of activation spanned the precuneus, bilateral angular gyri (AG), left inferior parietal lobe (IPL) and down the cingulum. In addition, there were small clusters of increased activity in the left insula and right supramarginal gyrus (SMG). In the known > unknown contrast, there was a cluster spanning the left and right precuneus, bilateral superior parietal lobe (SPL) and spreading to the left IPL. There were additional small clusters in the left middle temporal gyrus (MTG), left SMG and bilateral supplementary motor areas (SMA).

P2

In the trained > unknown contrast, there was a large cluster of activity encompassing the left IPL, AG and areas of the SPL. There was a further cluster including bilateral precuneus, right cuneus and right AG. In the known > unknown contrast there were

significant clusters of activity in bilateral SMG, bilateral precuneus, bilateral AG, areas of bilateral MTG and the right hippocampus spanning to right insula.

P3

In the trained > unknown contrast, there was a large cluster of activity spanning the right hippocampus, left thalamus spreading to the right inferior frontal gyrus (IFG) and mid frontal areas. There were additional clusters in the right AG/precuneus and the cerebellum. In the known > unknown contrast, there was a swathe of increased activity over the right postcentral gyrus and SMA spreading to the right STG, hippocampus, thalamus and precuneus. There were further clusters of increased activity in the left postcentral gyrus, left mid/inferior temporal gyrus, left IPL and the cerebellum.

P4

In the trained > unknown contrast, areas of increased activation included bilateral IFG, right anterior cingulum, anterior right SMA, left precentral gyrus, left putamen and insula. In the known > unknown contrast, there were similar areas of activation. Increased activation was observed in bilateral anterior cingulum spreading to the right SMA and areas of superior medial frontal cortex. There were clusters of activity in the bilateral IFG triangularis, the left insula spreading to the putamen, the left precentral gyrus and additional small clusters of activity in right midfrontal regions.

Table 3. Clusters of significant activity for naming of newly trained and untrained items.

Contrast	PWA	Peak region	Cluster size	Peak MNI			t
				x	y	z	
Trained > Untrained	P1	L MTG	127644	-48	-44	9	7.83
		L SMG		-61	-30	27	7.03
		L IPL		-55	-47	45	7.01
		R cerebellum	14940	25	-84	-30	5.05
		R cerebellum		44	-78	-32	5.01
		R AG	3617	45	-65	36	4.69
		R MTG		49	-52	15	3.69
	P2	R precuneus	12384	4	-73	43	5.38
		L precuneus		-5	-72	35	4.98
		L AG	10436	-37	-63	48	6.21
		L IPL		-34	-59	49	6.21
		R AG	2507	42	-54	38	4.77
		R IPL		38	-58	46	4.01
	P3	R fusiform	19222	30	-64	1	7.04
		L thalamus		-2	-18	13	6.47
		R thalamus		14	-33	10	4.89
		R hippocampus	4569	28	-22	-12	5.30
	P4	R mid cingulate	13514	4	-2	39	5.99
		L IFG op.	1131	-42	6	25	5.82
		L IFG tri.		-48	32	22	4.74
		R MFG	792	43	32	32	4.86
		R IFG tri.		44	27	26	4.51
		L putamen	45	-23	2	7	4.75
		L insula		-31	-3	15	3.99

Note: Clusters significant at $p < .001$ voxel height and $p < .05$ FWE cluster correction for picture naming of trained > untrained items. Up to three strongest peaks listed per cluster. L; left, R; right, MTG; middle temporal gyrus, SMG; supramarginal gyrus, IPL; inferior parietal lobe, AG; angular gyrus, IFG; inferior frontal gyrus, op; pars opercularis, tri; pars triangularis, MFG; middle frontal gyrus.

Table 4. Clusters of significant activity for naming of already known versus unknown items.

Contrast	PWA	Peak region	Cluster size	Peak MNI			t
				x	y	z	
Known > Untrained	P1	L IPL	23752	-55	-47	45	8.22
		L SMG		-61	-36	31	6.06
		L MTG	7006	-54	-55	12	5.49
	P2	L cuneus	18790	0	-92	25	5.73
		L lingual		12	-92	1	5.71
		R lingual		18	-77	8	4.40
		R MCC	7153	3	-40	52	6.36
		L precuneus		-5	-50	61	5.54
		L SPL		-21	-56	55	3.51
		R SMG	10994	62	-29	30	7.66
		R Rolandic		55	-16	24	6.63
		R STG		55	-34	15	3.30
		L postcentral	7669	-59	-23	32	6.59
		L SMG		-57	-33	40	6.40
		R IFG	2547	54	41	-8	5.05
	P3	R hippocampus	114602	27	-29	-6	6.29
		L lingual		-28	-62	3	6.11
		R thalamus		10	-28	11	5.66
		L ITG	4367	-57	-56	-9	4.40
		L MTG		-59	-62	0	4.32
	P4	L IFG	6328	-43	4	24	6.65
		L ACC		-1	14	32	6.22
		L insula		-35	-4	13	5.47
		L caudate	2468	-18	11	19	5.87

Note: Clusters significant at $p < .001$ voxel height and $p < .05$ FWE cluster correction. Up to three strongest peaks listed per cluster. L; left, R; right, IPL; inferior parietal lobe, SMG; supramarginal gyrus, MTG; middle temporal gyrus, MCC; midcingulate cortex, SPL; superior parietal lobe; SMG, supramarginal gyrus; IFG, inferior frontal gyrus; ITG, inferior temporal gyrus.

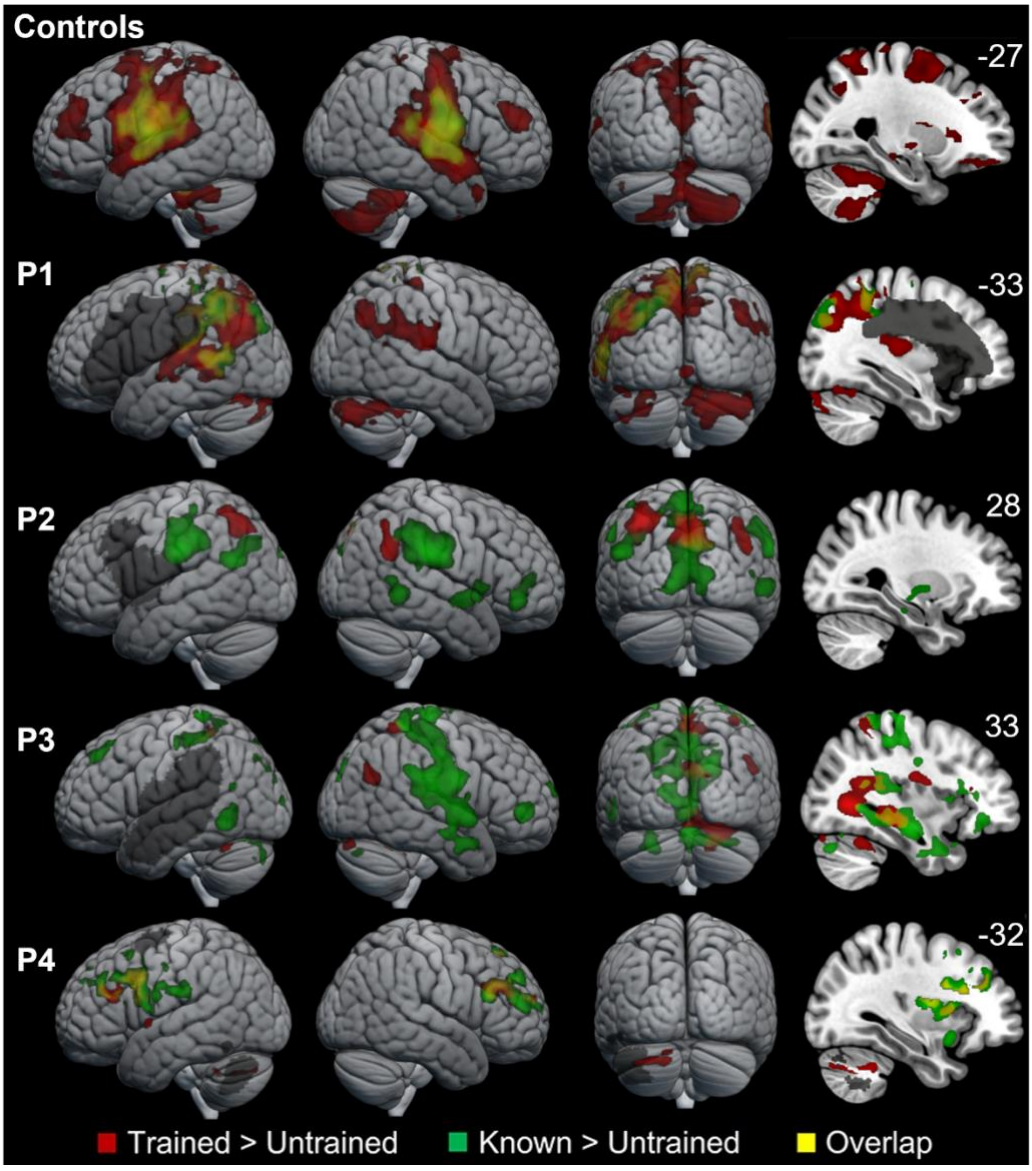


Figure 2. Whole brain activations for picture naming for controls and patients. Naming of newly trained > untrained (green) and known > untrained (red), with overlap (yellow). Top row: healthy matched controls (Chapter 1). Lesion profiles depicted in grey. Images thresholded at voxel level p < .001, cluster corrected FWE $p < .05$.

Anatomical connectivity mapping (ACM)

Each participant's anatomical connectivity map was contrasted with an average of the healthy control group ACMs. The contrast of PWA < healthy controls showed clusters of reduced structural connectivity in each participant. P1 had a swathe of reduced ACM spanning areas in the left inferior longitudinal fasciculus (ILF), arcuate fasciculus and inferior fronto-occiciptal fasciculus (IFOF; Figure 3a). P2 had areas of reduced ACM in the anterior segment of the left arcuate, IFOF and uncinate (Figure 3b). P3 had a cluster of reduced ACM in the left ILF (Figure 3c). P4 had no significant areas of reduced ACM (Figure 3d). There were no significant areas of increased ACM in any PWA over healthy controls.

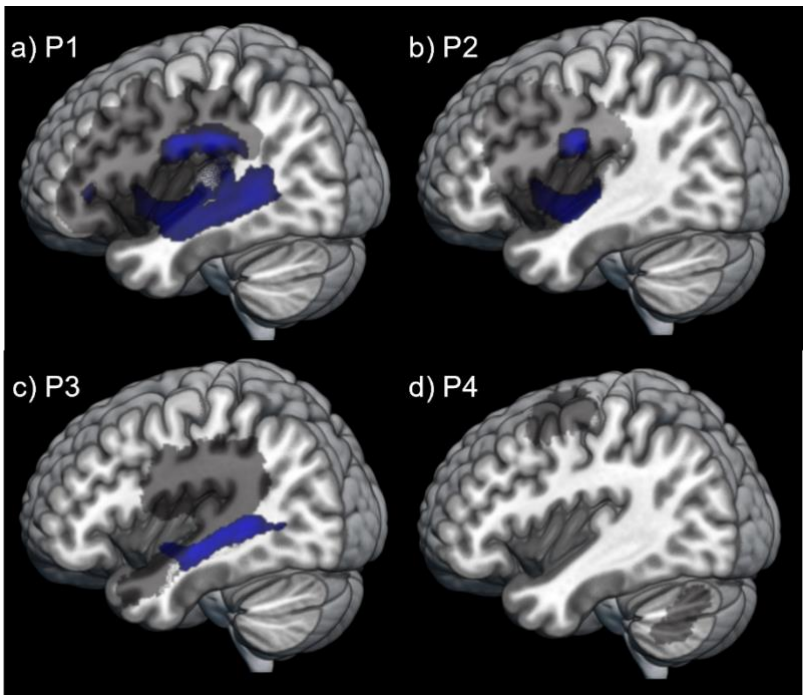


Figure 3. Areas of reduced structural connectivity in PWA compared to healthy controls. Lesion profiles (grey) and areas of reduced structural connectivity (blue). Images thresholded at voxel level .001, cluster corrected FWE $p < .05$.

DISCUSSION

The present study aimed to examine the neural mechanisms underlying novel word learning ability in post-stroke aphasia. PWA learned up to 50 novel, native words over the course of three weeks. PWA responded to completely unknown/untrained items, named successfully trained items and named previously known, and thus well consolidated, items during fMRI. Previously, control participants undertook the same procedure as used in this study to evaluate the role of the CLS framework in the acquisition of novel real-world vocabulary (Gore et al. 2021; Chapter 1). The results of the control participants and PWA were compared to evaluate the neural correlates of word learning in post-stroke aphasia.

PWA were able to acquire and maintain novel words successfully, with only a small non-significant disadvantage relative to healthy age-matched controls. Hence, it seems that language learning ability is intact in these PWA. In terms of the mechanisms underpinning this learning in PWA, these too seemed comparable to those seen in healthy age matched controls by Gore et al. (2021; Chapter 1), in that they fit with the predictions of the CLS model. When naming newly trained words, areas of the hippocampal-episodic memory network were activated, whereas when naming already known (and hence well consolidated) words, areas of the cortical language network were activated. Although all PWA successfully learned a similar number of items there was variability in level of learned item maintenance, and reduced maintenance of learning was associated with reduced connectivity of the AF and IFOF. Below, we consider the implications of each of these key findings for our understanding of the basis for novel word learning in chronic post-stroke aphasia.

Vocabulary acquisition

All four PWA successfully learned around 42 novel words over the course of 3 weeks training, on par with healthy control participants (Gore et al., 2021). Thus, despite significant left-hemispheric lesions PWA were able to effectively learn new novel items. However, the PWA in this study were well recovered with minimal language deficits. Consistent with previous literature, there was slight variation in individuals' performance, but all PWA reached high levels of learning. It is important to note there was no significant change in performance during pre-training baseline testing. Repeated testing alone can be sufficient to improve performance

(Howard et al. 2015), but as all trained items were previously unknown to participants the observed treatment gain is necessarily due to the training alone. Overall, these findings concur with previous literature, that PWA can learn novel words successfully (Gupta et al. 2006; Tuomiranta et al. 2012; Tuomiranta et al. 2015; Peñaloza et al. 2016).

Gore et al. (2021; Chapter 1) found that the acquisition of newly trained items by healthy age-matched controls was supported by a combination of regions associated with episodic memory (left hippocampus and left IPL) and language-semantic areas (left IFG and left ventral anterior temporal lobe). There was a shifting division of labour across these two memory networks in line with item consolidation status; faster naming was associated with more activation of language-semantic areas and lesser episodic memory activation. The finding of a neocortical shift associated with consolidation fits with the predictions of the CLS memory model. The question addressed in this study is to what extent such a model holds true under conditions of damage to the neocortical regions involved in language processing.

There were differing patterns of neural activity during naming of trained items between participants. The results indicate that PWA with larger lesions (P1, P2 and P3) had associated areas of consistent activation across the group when naming newly trained words, including the right AG. Although the most ventral left AG is associated with some semantic tasks like reading (Branzi et al. 2021), the core bilateral AG activity is likely to relate to episodic memory function (Humphreys et al. 2020). Further, amongst the PWA with large lesions the precuneus was consistently activated when naming trained words. It has been proposed that the precuneus may be casually involved in episodic memory retrieval (Wagner et al. 2005; Dörfel et al. 2009; Bonnì et al. 2015).

Across the three PWA with large lesions, the left posterior MTG (pMTG) was consistently activated. The pMTG is involved in verbal semantic cognition (Hoffman et al. 2012; Jackson 2021) and frequently damaged in patients with comprehension impairments post-stroke (Noonan et al. 2010). In addition, when naming already known items P3 had increased activity in the right IFG and right anterior temporal lobe (ATL). The left IFG is considered a critical region for speech production (Blank et al. 2002; Price 2011). There is considerable convergent evidence that the left

ventral ATL supports semantic memory (Lambon Ralph 2014; Lambon Ralph et al. 2017). Right hemispheric homologues of language regions such as the left IFG have been proposed as a compensatory mechanism of language recovery in aphasia (Winhuisen et al. 2005; Thiel et al. 2006). Therefore, these right-hemispheric activations may reflect language processing.

These results indicate that participants with larger lesions (P1, P2 and P3) relied upon a mixture of episodic and language regions to name new items. Thus, they could be considered in an intermediary stage of word learning as seen previously in healthy age-matched controls (Gore et al. 2021) and in PWA (Tuomiranta et al. 2015). This notion is supported by the three participant's maintenance performance, which showed significant drop-off, demonstrating many newly learnt items were not well consolidated, presumably due to damage to neocortical language regions. Interestingly, unlike healthy controls, PWA with larger lesions also relied upon episodic and language regions to name already known items. Already known items had a low age of acquisition and were consistently correctly named in baseline assessments. P2 and P3 had increased levels of left hippocampal activity when naming already known items. The hippocampus is considered the critical region supporting episodic memory (Eichenbaum et al. 1996; Tulving and Markowitsch 1998).

P4, who had a smaller, focal SMA lesion relied more upon language-semantic regions for naming both already known and newly trained words. This could indicate greater consolidation for P4 due to lesser damage to linguistic neocortex. Although P4 has an atypical and comparatively small lesion, it is notable that they had suffered severe chronic aphasia prior to recovery over the subacute stage and experienced persistent word finding difficulties. It would seem therefore that although P4's small left SMA lesion was sufficiently to produce ongoing production deficits, it nevertheless permitted unperturbed consolidation of new knowledge.

The results indicate that, like controls, PWA with larger lesions relied upon both episodic and language regions for naming of newly trained words. These results fit within the CLS model of memory, in a gradual transition from episodic memory processing to consolidated neocortical processing. This suggests that word learning in these PWA does not substantially differ from neurotypical word learning.

However, the baseline results of naming already known words are more difficult to interpret. Hippocampal activity is not typically associated with well consolidated picture naming. However, for all participants, accuracy for known items was not at ceiling during scanning and thus PWA were experiencing word finding difficulties. Hippocampal activation could indicate supplemental processes of episode searching as a compensatory strategy. Alternatively, methodological considerations could be attributed to the episodic processing demonstrated. As already known items were presented three times prior to training, participants could be recalling a specific episode of naming the identical picture stimuli. To avoid this potential confound, future studies could use multiple variants of the same target item, for example in different orientations or contexts.

Retention

Despite consistently high initial levels of acquisition across all PWA, there was nevertheless considerable variation in maintenance of trained items. At the top end of the scale, P4 retained every item over the course of three months. This is beyond the level of maintenance seen in the control participants and suggests a high level of consolidation of the novel items. P4 had a smaller, focal SMA lesion and these high levels of maintenance may indicate greater consolidation due to lesser damage in the language network. P3's maintenance was on par with controls, whereas P2 and P1 retained a smaller percentage of items than controls. Given the high levels of acquisition, this level of variability in retention is intriguing and suggest that consolidatory processes of word learning were disrupted for P1 and P2.

There was also variability in areas of reduced structural connectivity between participants. P1 and P2 had reduced structural connectivity in the AF and IFOF compared to age-matched controls. The AF is considered a key tract in language processing, associated with deficits in repetition (Berthier et al. 2012) and specifically successful novel word learning in healthy controls (López-Barroso et al. 2013). The left IFOF is suggested to be involved in the ventral pathway of language (Saur et al. 2008; Turken and Dronkers 2011). Both patients had good learning but reduced retention. Whereas P3, with reduced connectivity with the ILF, demonstrated good learning. These results indicate that spared connectivity in a combination of AF and IFOF was not necessary for acquisition but were necessary

for retention. Although the present results cannot distinguish which of these tracts are crucial for retention, it is possible that it is the combined action of the dorsal and ventral language tracts that act to bind the phonological form to the associated meaning and permit consolidation.

Methodological considerations

In treatment-based studies of aphasia there are two approaches to item selection for impairment directed therapy. A constant set of items for all patients allows for ease of administration and consideration, thus is used in many treatment studies (Abel et al. 2010, 2015; Nardo et al. 2017; Delikishkina et al. 2020). Whereas an individually tailored set of items that are not currently retrievable is the usual approach either explicitly or implicitly in clinical SLT (Palmer et al. 2017, 2019). In the former approach, milder patients would only be trained on higher frequency items that they can name already. In clinic, this leads to boredom effects and dropout (Brady et al. 2016). Using novel items, or individually tailored item sets in research may be beneficial for patient engagement, and thus more powerful results.

Conclusions

This study found that the neural underpinnings of novel word learning in post-stroke aphasia mirrored that of healthy controls, fitting within the CLS model. Naming of newly learned items was associated with increased activation of both episodic and language regions. However, unlike controls, naming of already known items was also associated with episodic and language region activation in PWA with larger lesions. Additionally, PWA maintained a variable amount of learned item names. Participants with reduced structural connectivity in the AF and IFOF retained fewer items over a period of three months. These results make further our understanding of knowledge acquisition under conditions of neocortical damage through demonstrating that although neural function underpinning language acquisition in these PWA was largely normal, there were structural abnormalities in the dorsal and ventral pathways associated with the longevity of retention.

CHAPTER 6: GENERAL DISCUSSION

This final Chapter is split into four sections. In the first section, the key findings from each of the five empirical Chapters are summarised. The second section discusses the implications of these findings in the context of neurotypical vocabulary acquisition and the broader theoretical implications. Next, the clinical and theoretical implications of these findings in aphasia are discussed, including reflections upon how these findings could influence aphasiological treatment. In the fourth section, consideration is given to potential directions for future research.

The overarching aim of this thesis was to enhance our understanding of the neural correlates underlying word learning processing, integrating evidence from healthy older adults and individuals with post-stroke aphasia. A cross-cutting theme within this thesis is the Complementary Learning Systems (CLS; Marr, 1971; McClelland, McNaughton, & O'Reilly, 1995) model. This theory proposes that new knowledge is initially coded through rapidly established, sparse representations of episodes supported by episodic processing regions, including the hippocampus. Slower, deeper learning in domain-specific neocortical regions supports longer-term consolidation and the development of generalisable representations across episodes. Over time, there is a gradual shift in the division of labour between MTL and neocortical regions (with varying rate of change: cf. McClelland et al., 2020).

An emergent theme within this thesis was the set of ideas termed "variable neuro-displacement" (Stefaniak et al. 2020). This set of ideas comes from a fundamental principle of engineering that well-engineered systems are resilient to functional stresses whilst balancing performance and energy costs. Recovery from post-stroke aphasia may involve engagement of degenerate networks or spare capacity within networks. Neural processing is metabolically costly. Thus, under standard levels of performance demand, this spare capacity may be downregulated to reduce energy usage. When task demand increases or network efficiency is reduced via damage, activity in these areas may be upregulated (Stefaniak et al. 2020). These ideas were explored in healthy word learning.

The specific research objectives outlined in the Introduction were as follows.

(i) To investigate whether word learning in healthy populations fits within both stages of the CLS model; initial learning supported by hippocampal-episodic memory systems, followed by a gradual shift in the division of labour to domain-specific neocortical languages areas.

(ii) Explore the levels of interactivity between episodic and language-semantic systems at differing stages of consolidation.

(iii) Examine whether the set of ideas termed variable neuro-displacement are applicable beyond recovery in post-stroke aphasia to neurotypical language learning.

(iv) Does re-learning or novel word learning in post-stroke aphasia fit within this healthy learning framework? If so, is there a level of critical damage to the dominant left hemispheric language regions, resulting in deviations from the norm?

SUMMARY OF FINDINGS

In Chapter 1, 20 healthy older participants were trained on novel, native words for three weeks with post-learning functional neuroimaging. Newly learned items were compared to two conditions: (i) previously known items to highlight the similarities and differences with established vocabulary, and (ii) unknown/untrained items to control for non-specific perceptual and motor-speech output. The findings of this Chapter map a framework for novel, native word learning in healthy older adults. These results provide neural evidence of CLS in vocabulary acquisition. Naming newly trained items was associated with increased hippocampal activity at a whole-brain level. This greater left hippocampal activity was associated with worse behavioural performance (reduced accuracy, longer RTs and less trained item name retention).

In contrast, there was no association between hippocampal activity and performance when naming already known items. The second consolidation stage of the CLS proposes a gradual shift from reliance on the MTL-episodic network towards long-term neocortical consolidation. In line with this prediction, we found that when

naming newly trained words, higher levels of left IFG and ATL activation were associated with better accuracy and shorter RTs. Crucially, the associations in each region of interest between BOLD activity and performance were significantly different between naming of previously known items and newly trained items. Overall, the findings in Chapter 1 provide evidence for both aspects of the CLS model in long-term, native word acquisition.

Although Chapter 1 and previous literature provide evidence of CLS in word acquisition, the question remained as to how these memory systems interact during varying stages of consolidation. In Chapter 2, this proposed neural shifting division of labour between complementary systems was investigated in the healthy control data from Chapter 1. Similarly, naming already known, newly learned, and unknown/untreated words provided a snapshot of consolidation across word learning. PPI analyses were performed to assess functional connectivity between a critical neocortical language region (IFG) and the rest of the brain at different stages of consolidation. Regardless of behavioural performance, there was increased functional connectivity between left IFG and the picture naming network when naming well-consolidated, already known words. Conversely, when naming newly learned words, there was increased connectivity between the left IFG and the same naming network, plus areas involved in effortful episodic retrieval (left hippocampus and bilateral intraparietal sulcus) associated with decreased behavioural performance. These results suggest that there is a time-limited intermediate stage in which both MTL-episodic and neocortical-language processing support new knowledge.

In Chapter 3, the structural correlates of word learning were explored. Right hemispheric structural neural correlates corresponded with increased word learning ability. Right inferior frontal gyrus structural intensity measures correlated with better behavioural performance and less activation of episodic regions for naming of newly trained items. These associations were moderated by structural intensity of the left inferior frontal gyrus. Individuals with higher structural intensities in the dominant left hemispheric language regions did not manifest right-sided functional or structural associations. The right hemisphere was required structurally when there was decreased structural intensity in critical regions. This relationship was interpreted in the context of a set of ideas termed “variable neuro-displacement”, a

principle that has been previously applied to aphasia and may be applicable across many different populations.

In Chapter 4, a "reverse translation" study, 16 patients with chronic post-stroke aphasia undertook three weeks of speech and language therapy interventions. The stimuli in this study differed from the complex, rare items in Chapters 1-3. Instead, these items were everyday, high word frequency items drawn from the International Picture Naming Project. Patients completed a subsequent fMRI picture naming task, which probed (i) treated items, (ii) already known items and (iii) untreated/unknown items. Overall, the results demonstrated that patients relied on similar underlying neural mechanisms as healthy controls when re-learning items, fitting within the CLS framework. Increased levels of remaining left inferior frontal gyrus (IFG) activity was associated with quicker responses, and the level of damage to this critical language region moderated the relationship between hippocampal activity and the naming of treated items. In patients with relatively preserved left IFG, hippocampal activity was associated with lower accuracy and slower responses. However, patients with over 50-60% damage to the left IFG displayed the opposite pattern, with more hippocampal activity when naming associated with quicker responses. Hippocampally-dependent patients may be either (i) in the initial learning phase due to a lack of retained partial representations or (ii) unable to shift the division of labour to the damaged cortical language systems.

In Chapter 5, the functional neural correlates of novel word learning in post-stroke aphasia were explored. In this case series, patients learned the same complex, rare items as in Chapters 1-3. Patients successfully learned these difficult novel items on a par with healthy controls. Also similar to controls, naming of newly learned items was associated with increased activation of both episodic and language regions. However, the PWA differed in neural activity when naming already known items. Patients with larger lesions had episodic and language region activation during naming of items which they could name prior to training. Additionally, participants with reduced structural connectivity in the AF and IFOF retained fewer items over a period of three months. Although the functional correlates underpinning language acquisition in these PWA was largely normal, there were structural abnormalities in the dorsal and ventral pathways associated with the longevity of retention.

VOCABULARY ACQUISITION

The Chapters in this thesis contribute to our understanding of neurotypical novel vocabulary acquisition. Convergent evidence of complementary memory systems in novel, native word learning was demonstrated in univariate functional, structural and connectivity analyses. These results align with and extend previous literature supporting Davis and Gaskell's (2009) proposal of the CLS framework in vocabulary acquisition. These multimodal results provide insight into the stages of consolidation within word learning (Chapter 1), the timescale of these stages (Chapter 2) and what other variables can affect the timescales of these stages (Chapter 3).

CLS in vocabulary acquisition

There are multiple previous sources of evidence for the CLS theory in word learning. In terms of the initial stage in the CLS model, there is converging evidence of hippocampal involvement in various aspects of word learning. In repetition tasks, mapping of auditory representations of speech (hearing the stimuli) and motor representations of speech (repeating the items) are required. Pseudoword learning, with an associated repetition task, has been associated with hippocampal involvement in the initial stages of word learning, with decreased activation over repeated trials (Davis et al. 2009). Pseudoword learning through silent sentence reading has also been associated with MTL activation (Mestres-Missé et al. 2008). This process requires linking of visual orthographic form with context-specific meaning.

The results of this thesis align with and expand upon these previous studies with the addition of spoken item name recall. In Chapters 1-3, healthy participants learned item name orthography, phonology and associated semantic knowledge. There were significant increases in behavioural performance for verbal naming and semantic knowledge of the items. Better learning of semantic information about the novel items corresponded with better naming performance, a finding which has been previously demonstrated (Takashima et al. 2017).

In these Chapters, the neural correlates of naming unknown, newly-learned and already known items were compared. The clusters of activation for naming of

already known items provided a clear baseline of regions supporting fully consolidated words. Learning tasks often suffer from a lack of knowledge as to what regions would support the task post-learning. Although this may be countered somewhat with pre- and post-learning scanning sessions, there is evidence that learning of words is a long process. Thus, post-learning imaging may be indicative of an intermediary stage of consolidation rather than well consolidated, entirely domain-specific processing.

Naming of newly learned items was supported by a combination of episodic and language processing regions, whereas naming of well-consolidated already known items was associated with language processing areas (Chapter 1). These Chapters provide evidence of the first stage of the CLS in: (i) lexical-semantic integration, learning verbal item names for pictures, orthographic word forms and semantic knowledge; (ii) auditory-motor integration, mapping of sound to articulatory action to produce name the pictures overtly. (iii) episodic-lexical integration, the consolidation of initial short-term, episodic representations of novel words to deeper, long-term representations in neocortical language areas.

Timescales in the shifting division of labour

In the CLS model, hippocampal reinstatement of activation patterns for episode specifics is proposed as a mechanism of consolidation (McClelland et al., 1995) and hippocampal replay during sleep (Wilson and McNaughton 1994). Sleep has also been suggested as a means for instantiating hippocampal representations of novel words for neocortical learning (Davis and Gaskell 2009). Sleep-dependent word knowledge consolidation is demonstrated in several studies with slow-wave sleep associated with improved recognition of new knowledge (Tamminen et al. 2010, 2013; Lewis and Durrant 2011; Gaskell et al. 2019).

The empirical studies within this thesis were not designed to isolate the specific effects of sleep on word learning. Instead, the learning stage was purposely long to allow for sufficient sleep periods. Thus, there was a more extended period for repeated exposure, sleep and associated decreases in hippocampal activation. Nevertheless, when naming newly trained items, there was still significant hippocampal involvement, indexed by level of consolidation. Previous neuroimaging studies of novel word learning provide an intriguingly quick transition of labour,

with hippocampal activation declining over presentations within a single scanning session (Breitenstein et al. 2005; Davis et al. 2009) or over 24 hours (Davis et al. 2009). However, there is also neuropsychological evidence that although vocabulary acquisition processes are initialised relatively quickly, a shift in labour division may depend on a critical level of exposure and learning (Gaskell and Dumay 2003). The learning task in this thesis was substantial, with phonologically complex item names in overlapping categories of items.

The results of these healthy word learning Chapters converge, alongside previous literature, demonstrate; (i) an intermediate stage of consolidation in word learning, with a split in the division of labour between episodic and language-semantic regions when naming newly trained items; (ii) functional connectivity between episodic and language-semantic regions during this intermediary stage, indexed by level of consolidation; (iii) a 'final' stage of consolidation, with only language-semantic region processing when naming well consolidated already known items. Taking these results together with the body of previous literature, there is clearer evidence of the CLS model at three key stages of word learning.

Variable neuro-displacement in novel word learning

The idea of variable neuro-displacement aligns with research on post-stroke aphasia recovery (Chang and Lambon Ralph 2020; Stefaniak et al. 2020, 2021). However, it was previously unclear whether these ideas could be applied beyond aphasia as a general principle. In Chapter 2, these ideas were extended to normal language learning. Increased structural intensity in regions of the right hemisphere was associated with better word learning performance. However, these effects were only seen for individuals with decreased structural intensities in dominant left hemispheric language regions.

These results align with a computational model of spoken language production with bilateral pathways (Chang and Lambon Ralph 2020). An imbalance in processing capacity was applied to the otherwise homogenous bilateral structure, whereby there were more hidden units in the left hemisphere. This bilateral architecture and capacity imbalance was sufficient to model patterns of neural data demonstrated in both healthy participants and patients with post-stroke aphasia. In the data of this thesis, individuals with sufficient structural intensity and, therefore, capacity in the

left hemisphere did not require right hemispheric input. In contrast, participants with weaker structural intensities in the dominant left hemisphere were reliant upon bilateral processing, with right hemispheric areas of the language network picking up some performance demands.

VOCABULARY RE-ACQUISITION

CLS theory in PWA word re-learning

Brain-behaviour correlations of word re-learning in PWA followed a similar pattern as seen in healthy word learning (Chapter 1) up to a critical point of damage. These results are important theoretically, as several longstanding theories of aphasia recovery require a great deal of neuroplasticity, with regions premorbidly not involved in language tasks taking on new roles. A more biologically plausible explanation of recovery in aphasia is that neural systems are biologically resilient to increased performance demands and damage (Stefaniak et al. 2020). Patients with sufficiently spared regions of neocortical language processing had enough spare capacity within the network to pick up the additional demands caused by damage to the network. Patients with more extensive damage to dominant left hemispheric language regions did not have enough capacity within the network. Although the language network may be architecturally bilateral, the dominant left hemisphere may be more efficient with spare capacity. The right hemisphere may have some capacity for processing, but this was not sufficient to pick up demands with more extensive left hemispheric lesions. Thus, these patients were reliant upon alternate networks, such as episodic memory systems and the executive functioning network.

As previously discussed, bilateral networks may provide resilience to damage. These results also align with the bilateral computational model of speech production (Chang and Lambon Ralph 2020). In this model, the left pathway has greater processing capacity. When this capacity imbalance is combined with damage, there is evidence of recovery via a reoptimisation of connection weights, analogous to plasticity-related recovery in post-stroke aphasia. With minimal damage to connection weights, the left hemisphere dominates activation of units. However,

with damage to right pathway connections, there is insufficient capacity with the right pathway to compensate completely.

Subject-level versus group-level inferences

Although word re-learning in post-stroke aphasia was comparable to the healthy control framework at region of interest level, the group whole brain activations in Chapter 4 were not as similar. The clusters of activation for group picture naming were in a reasonable location, as seen previously in studies of PWA speech production (Stefaniak et al. 2021) and controls for the same task (Chapter 1). However, these clusters were of a small extent and constrained to the right hemisphere, whereas speech production in controls is associated with bilateral activation of the pSTG and areas of the left prefrontal cortex (Blank et al. 2002; Price 2010, 2012). Patients with post-stroke aphasia have such graded variability over so many factors that significant effects may be cancelled out.

It is important to consider the implications of a group level analysis in aphasia. Group studies provide the opportunity to make wider inferences beyond the individual, validating treatment methods and exploring options for stratification of patients (Howard 2003). However, the level of inhomogeneity within the group due to lesion profiles, premorbid language ability and post-damage cognitive deficits can result in highly variable patterns of behavioural improvement and changes in functional activation (Howard 2003). ROI analyses avoid the issues of averaging over a heterogenous group.

However, these functional-behavioural interactions may still be influenced by variation in lesion profiles. In Chapter 4, there were opposite patterns of functional-behavioural correlations at different levels of damage to the left IFG. These results indicate the important of considering lesion profiles within group aphasia studies. An alternative is to have a much larger sample size, with associated increases in power. Generally, averaging over large groups of individuals is beneficial to smooth effects of individual differences and garner insight over the general population of the sample (Grady et al. 2021). However, there were between-subject differences in the case series of Chapter 5. These small nuances may be missed in group studies (Lorca-Puls et al. 2018).

As large group studies become the gold standard in neuroimaging literature, it is important not to lose sight of the value of single-case and case-series analyses. Subject-level inference allows detailed specification of the participant, treatment, and outcomes of this treatment. Single-case studies can also be analysed in native space, avoiding issues of inter-subject registration and allowing for correspondence between individual patient anatomy and activation. Subject-level inference may therefore also increase the likelihood of visualising perilesional activation (Howard 2003). However, one of the main disadvantages of subject-level inference is that the results are true for only one participant and do not necessarily generalise to other participants. A combination of larger group studies and single-case or case-series designs may provide the necessary power and specificity to detect graded effects.

Clinical implications

The treatments used in Chapters 4 and 5 of this thesis were successful at a group and individual subject level. Although there was variability amongst the PWA as to therapy gains, overall, there were significant treatment outcomes. PWA with relatively mild deficits in Chapter 5 could learn novel, native words with performance levels similar to healthy controls. For these individuals, the standard items used in the group study of Chapter 4 would not be suitable and very quickly induce boredom effects. In treatment-based studies of aphasia there are two approaches to item selection for impairment directed therapy. A constant set of items for all patients allows for ease of administration and consideration, thus is used in many treatment studies (Abel et al. 2010, 2015; Nardo et al. 2017; Delikishkina et al. 2020). Whereas an individually tailored set of items that are not currently retrievable is the usual approach either explicitly or implicitly in clinical SLT (Palmer et al. 2017, 2019). In the former approach, milder patients would only be trained on higher frequency items that they can name already. In clinic, this leads to boredom effects and dropout (Brady et al. 2016). Using novel items, or individually tailored item sets in research may be beneficial for patient engagement, and thus more powerful results.

Individuals with more extensive damage to the language network may not have enough availability within the domain-specific neocortical memory systems available to allow for a shift in the division of labour to this damaged network. If

individuals cannot adequately consolidate this knowledge, they may remain reliant on short-term MTL-hippocampal episodic memory systems to use a learned/re-learned word. If these theories were to hold, it may be beneficial for patients with larger lesions not abruptly to stop treatment, as is often the case in SLT. Instead, lower intensity treatments could be continuously undertaken for long periods of time. This may be enough to bolster short-term memory systems, keeping a re-learned item available for everyday use.

Providing continuous SLT is not a feasible proposal. However, the patients in this thesis had generally successful treatment outcomes with a self-guided, computerised therapy programme. Computerised therapy does not require a therapist to be present and, therefore, has lower costs. These results align with previous literature, demonstrating the positive effects of computerised therapy (Palmer et al. 2019). Thus, perhaps self-guided therapy could aid in the prolonged treatment of patients with larger lesions who are potentially hippocampally-dependent.

These results may suggest that patients with larger lesions could benefit from alternative forms of therapy. Phonological short-term memory ability correlates with the ability to acquire vocabulary in PWA and healthy participants (Gupta et al. 2006). Few verbal short-term memory treatment studies have been conducted in aphasia, however there is initial evidence of positive outcomes in some patients (Minkina et al. 2017). There was evidence of executive and attention network activity supporting word learning in Chapter 5, perhaps because of variable neuro-displacement. If these networks are important in variable neuro-displacement and thus recovery, PWA may benefit from more engaging treatment. Therapies such as RISP (Conroy et al. 2018) used in this thesis use a game-like structure, where PWA attempt to beat the computer at naming. Further gamification of therapies may drive engagement and subsequently increase executive system involvement in language function.

FUTURE DIRECTIONS

The results described in this thesis raise further theoretical questions. Potential options for future directions using the data from this thesis and further empirical investigations to explore these hypotheses are described below.

The consolidation process could be considered incomplete because participants still used a combination of episodic and semantic processing when naming the learned items. An alternative approach to exploring the timescales involved in the gradual shifting of labour could involve a more extended learning period. However, the necessary length of training is difficult to determine. Initial scanning pre-learning, scanning during learning and subsequent post-scanning may provide a clearer picture of timescales in the functional connectivity changes supporting consolidation.

Computational modelling

Another potential direction of study to explore these questions is computational modelling. Computational models of knowledge acquisition often focus on either an episodic or semantic storage area, using interleaved learning to avoid catastrophic interference. To enable disentanglement of functional connectivity between episodic and semantic systems, a model is needed that includes both processes. A potential option is a fully connected recurrent model. The model would learn both a) an episodic task such as veridical recall of a timestamped input vector from a series of inputs and b) a semantic task; requiring generalisation across multiple episodic vectors to identify the underlying semantic structure.

The key test of such a model would be the ability to both generate episodic veridical recall and learn the underlying semantic structure without catastrophic interference of learned concepts. Further, would this ability vary depending on the direction and extent of the connections between the episodic and semantic regions? Simulations such as these could be compared to neuroimaging results of healthy word learning, using output unit activation as a proxy for BOLD activation and functional connectivity measures. The pathways between input and episodic/semantic layers could be lesioned individually, adding noise to the input to simulate hippocampal amnesia and semantic dementia, respectively. Systematically damaging models in

this way may build upon the results of this thesis to enable more mechanistic explanations of language recovery in post-stroke aphasia.

CLS theory in semantic learning

There were relatively sparse results in the ventral anterior temporal lobes (vATL) in these Chapters. The vATL has been proposed as a trans-modal hub key to semantic representation (Lambon Ralph et al. 2016). There is converging evidence for this proposal from neuropsychological studies (Warrington 1975; Jefferies and Lambon Ralph 2006b; Patterson et al. 2007), computational modelling (Rogers et al. 2004; Chen et al. 2017; Jackson et al. 2021), neuroimaging (Binney et al. 2010; Visser et al. 2012; Rice et al. 2018) and transcranial magnetic stimulation (Pobric et al. 2007). However, the vATL are susceptible to signal loss in functional imaging due to local differences in magnetic susceptibility at the interface of tissue and fluid/air (Devlin et al. 2000). To increase the likelihood of reducing signal loss in the temporal lobes, a field of view of more than 15cm, a 45-degree tilt from the AC-PC and a higher-level baseline were employed suggested in Visser, Jefferies, & Lambon Ralph (2010). In addition, dual-echo paradigms have been previously shown to overcome temporal lobe signal loss by optimally combining echoes (Poser et al. 2006; Halai et al. 2014). Thus, a multi-echo paradigm was employed.

The inclusion of the vATL as a region of interest can increase power. There were correlations between the vATL and behavioural performance when an ROI was employed in Chapter 1. However, there was a lack of activation in the bilateral temporal lobes at a whole brain level during the picture naming tasks. Overt speech production was required in these tasks, which inevitably introduces motion-induced signal (Fiez 2001). Although participants were asked to speak like a ventriloquist and spent some time prior to scanning practising this, there were some motion artefacts in the data. Using three echoes at first inspection may seem illogical, as the third echo is of a sub-optimal length, three echoes allow for echo time (TE) relaxation curve estimation. Blood oxygenation level-dependent (BOLD) signal is dependent on the TE relaxation curve, whereas noise is not. This estimation allows for the automatic or manual removal of non-BOLD signals and can mitigate motion artefacts beyond standard removal (Gonzalez-Castillo et al. 2016) using multi-echo independent component analysis (ME-ICA; Kundu *et al.*, 2012, 2017) denoising.

Questions remain as to whether aspects of semantic learning fit within CLS theory. Future studies could employ an explicitly semantic learning paradigm to determine whether long-term acquisition of word meanings is initially supported via hippocampal processing. Alternatively, more naturalistic tasks could be used to isolate semantic learning. For example, while watching a television program series, there is generalisation across episodes to learn various relationships. These include relationships between characters, e.g., who belongs to which family; relationships between characters and location, e.g., where someone lives. Imaging semantic task processing pre-learning, during and post-learning may provide further information on the role of consolidatory processes in semantic learning.

CONCLUSIONS

The overarching finding of this thesis was that word learning in healthy controls and re-learning in post-stroke aphasia is supported by complementary episodic and neocortical language memory systems, aligning with the CLS model of knowledge acquisition. There was evidence of a gradual shift in the division of labour from episodic processing to domain-specific neocortical languages areas in healthy controls. Connectivity between the episodic and language networks was time-limited to this intermediary consolidation stage. Patients with mild damage and language deficit could learn and maintain novel items on par with healthy controls. The patterns of activation supporting these processes in patients with mild aphasia mirrored that of controls. However, after a critical level of damage to dominant language processing regions, the shifting division of labour was disrupted.

REFERENCES

Abel S, Schultz A, Radermacher I, Willmes K, Huber W. 2003. Increasing versus vanishing cues in naming therapy. 87:143–144.

Abel S, Schultz A, Radermacher I, Willmes K, Huber W. 2005. Decreasing and increasing cues in naming therapy for aphasia. Aphasiology. 19:831–848.

Abel S, Weiller C, Huber W, Willmes K, Specht K. 2015. Therapy-induced brain reorganization patterns in aphasia. Brain. 138:1097–1112.

Abel S, Willmes K, Huber W. 2007. Model-oriented naming therapy: Testing predictions of a connectionist model. Aphasiology. 21:411–447.

Acosta-Cabronero J, Patterson K, Fryer TD, Hodges JR, Pengas G, Williams GB, Nestor PJ. 2011. Atrophy, hypometabolism and white matter abnormalities in semantic dementia tell a coherent story. Brain. 134:2025–2035.

Anderson JM, Gilmore R, Roper S, Crosson B, Bauer RM, Nadeau S, Beversdorf DQ, Cibula J, Rogish M, Kortencamp S, Hughes JD, Gonzalez Rothi LJ, Heilman KM. 1999. Conduction aphasia and the arcuate fasciculus: A reexamination of the Wernicke-Geschwind model. Brain Lang. 70:1–12.

Andersson JLR, Skare S, Ashburner J. 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. Neuroimage. 20:870–888.

Andersson JLR, Sotiropoulos SN. 2016. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. Neuroimage. 125:1063–1078.

Axmacher N, Elger CE, Fell J. 2008. Ripples in the medial temporal lobe are relevant for human memory consolidation. Brain. 131:1806–1817.

Bayley PJ, Hopkins RO, Squire LR. 2006. The Fate of Old Memories after Medial Temporal Lobe Damage. J Neurosci. 26:13311.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B. 57:289–300.

Berthier ML. 2005. Poststroke Aphasia. Drugs Aging. 22:163–182.

Berthier ML, Lambon Ralph MA, Pujol J, Green C. 2012. Arcuate fasciculus variability and repetition: The left sometimes can be right. Cortex. 48:133–143.

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

Best W, Greenwood A, Grassly J, Herbert R, Hickin J, Howard D. 2013. Aphasia rehabilitation: Does generalisation from anomia therapy occur and is it predictable? A case series study. Cortex. 49:2345–2357.

Binder JR, Desai RH, Graves WW, Conant LL. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex. 19:2767–2796.

Binney RJ, Embleton K V, Jefferies E, Parker GJM, Lambon Ralph MA. 2010. The Ventral and Inferolateral Aspects of the Anterior Temporal Lobe Are Crucial in Semantic Memory: Evidence from a Novel Direct Comparison of Distortion-Corrected fMRI, rTMS, and Semantic Dementia. Cereb Cortex. 20:2728–2738.

Blank SC. 2002. Speech production: Wernicke, Broca and beyond. Brain. 125:1829–1838.

Blank SC, Scott SK, Murphy K, Warburton E, Wise RJS. 2002. Speech production: Wernicke, broca and beyond. Brain. 125:1829–1838.

Blasi V, Young AC, Tansy AP, Petersen SE, Snyder AZ, Corbetta M. 2002. Word retrieval learning modulates right frontal cortex in patients with left frontal damage. Neuron. 36:159–170.

Bonilha L, Gleichgerrcht E, Nesland T, Rorden C, Fridriksson J. 2016. Success of Anomia Treatment in Aphasia Is Associated With Preserved Architecture of Global and Left Temporal Lobe Structural Networks. Neurorehabil Neural Repair. 30:266–279.

Bonnì S, Veniero D, Mastropasqua C, Ponzo V, Caltagirone C, Bozzali M, Koch G. 2015. TMS evidence for a selective role of the precuneus in source memory retrieval. Behav Brain Res. 282:70–75.

Bookheimer SY, Zeffiro TA, Blaxton T, Gaillard W, Theodore W. 1995. Regional cerebral blood flow during object naming and word reading. Hum Brain Mapp. 3:93–106.

Bozeat S, Lambon Ralph MA, Patterson K, Garrard P, Hodges JR. 2000. Non-verbal semantic impairment in semantic dementia. Neuropsychologia. 38:1207–1215.

Bozic M, Tyler LK, Ives DT, Randall B, Marslen-Wilson WD. 2010. Bihemispheric foundations for human speech comprehension. Proc Natl Acad Sci U S A. 107:17439–17444.

Bozzali M, Parker GJM, Serra L, Embleton K, Gili T, Perri R, Caltagirone C, Cercignani M. 2011. Anatomical connectivity mapping: a new tool to assess brain disconnection in Alzheimer's disease. Neuroimage. 54:2045–2051.

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

Branzi FM, Pobric G, Jung J, Lambon Ralph MA. 2021. The Left Angular Gyrus Is Causally Involved in Context-dependent Integration and Associative Encoding during Narrative Reading. J Cogn Neurosci. 33:1082–1095.

Breitenstein C, Jansen A, Deppe M, Foerster A-F, Sommer J, Wolbers T, Knecht S. 2005. Hippocampus activity differentiates good from poor learners of a novel lexicon. Neuroimage. 25:958–968.

Brett M, Anton J-L, Valabregue R, Poline J-B. 2002. Region of interest analysis using an SPM toolbox [abstract] Presented at the 8th International Conference on Functional Mapping of the Human Brain. Sendai, Japan Available on CD-ROM in NeuroImage. 16.

Broca PP. 1861. Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). Publ **1861** Paris by Masson.

Buchsbaum BR, Olsen RK, Koch P, Berman KF. 2005. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. Neuron. 48:687–697.

Burgess PW, Shallice T. 1997. The Hayling and Brixton Tests: Thames Valley Test Company Bury St, Edmonds, England. Bury St Edmunds: Thames Valley Test

Company.

Burke DM, Shafto MA. 2008. Language and aging. In: Craik F. I. M.,, Salthouse T. A., editors. The handbook of aging and cognition. Psychology Press. p. 373–443.

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

Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. 2008. The parietal cortex and episodic memory: An attentional account. Nat Rev Neurosci.

Carr MF, Jadhav SP, Frank LM. 2011. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. Nat Neurosci 2011 142. 14:147–153.

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

Chang YN, Lambon Ralph MA. 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. 117:32779–32790.

Chen L, Lambon Ralph MA, Rogers TT. 2017. A unified model of human semantic knowledge and its disorders. Nat Hum Behav. 1.

Code C. 2003. The quantity of life for people with chronic aphasia. Neuropsychol Rehabil. 13:379–390.

Cohen L, Dehaene S, Naccache L, Lehéricy S, Dehaene-Lambertz G, Hénaff MA, Michel F. 2000. The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain. 123:291–307.

Community T tedana, Ahmed Z, Bandettini PA, Bottenhorn KL, Caballero-Gaudes C, Dowdle LT, DuPre E, Gonzalez-Castillo J, Handwerker D, Heunis S, Kundu P, Laird AR, Markello R, Markiewicz CJ, Maullin-Sapey T, Moia S, Salo T, Staden I, Teves J, Uruñuela E, Vaziri-Pashkam M, Whitaker K. 2021. ME-ICA/tedana: 0.0.11.

Conroy P, Sage K, Lambon Ralph MA. 2009. The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia. Aphasiology. 23:707–730.

Conroy P, Sotiropoulou Drosopoulou C, Humphreys GF, Halai AD, Lambon Ralph MA. 2018. Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia. Brain. 141:1815–1827.

Consortium B. 2007. The British National Corpus.

Coran M, Rodriguez-Fornells A, Ramos-Escobar N, Laine M, Martin N. 2020. Word Learning in Aphasia: Treatment Implications and Structural Connectivity Analyses. Top Lang Disord. 40:81–109.

Cornelissen K, Laine M, Renvall K, Saarinen T, Martin N, Salmelin R. 2004. Learning new names for new objects: Cortical effects as measured by magnetoencephalography. Brain Lang. 89:617–622.

Cox RW. 1996. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res. 29:162–173.

Cox RW, Hyde JS. 1997. Software tools for analysis and visualization of fMRI data. NMR Biomed. 10:171–178.

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

Crinion JT, Leff AP. 2007. Recovery and treatment of aphasia after stroke: functional imaging studies. Curr Opin Neurol. 20:667–673.

Davis MH. 2003. Connectionist modelling of lexical segmentation and vocabulary acquisition. In: Quinlan P, editor. Connectionist models of development: Developmental processes in real and artificial neural networks. Psychology Press. p. 151–187.

Davis MH, Di Betta AM, Macdonald MJE, Gaskell MG. 2009. Learning and consolidation of novel spoken words. J Cogn Neurosci. 21:803–820.

Davis MH, Gaskell MG. 2009. A complementary systems account of word learning:

neural and behavioural evidence. Philos Trans R Soc B Biol Sci. 364:3773–3800.

Davis MH, Johnsrude IS. 2007. Hearing speech sounds: Top-down influences on the interface between audition and speech perception. Hear Res. 229:132–147.

Dehaene S, Cohen L. 2011. The unique role of the visual word form area in reading. Trends Cogn Sci. 15:254–262.

Dehaene S, Cohen L, Sigman M, Vinckier F. 2005. The neural code for written words: A proposal. Trends Cogn Sci. 9:335–341.

Dehaene S, Le Clec'H G, Poline J-B, Le Bihan D, Cohen L. 2002. The visual word form area: a prelexical representation of visual words in the fusiform gyrus. Neuroreport. 13:321–325.

Delikishkina E, Lingnau A, Miceli G. 2020. Neural correlates of object and action naming practice. Cortex. 131:87–102.

Dell GS. 1986. A spreading-activation theory of retrieval in sentence production. Psychol Rev. 93:283–321.

Dell GS, Lawler EN, Harris HD, Gordon JK. 2004. Models of errors of omission in aphasic naming. Cogn Neuropsychol. 21:125–145.

Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. 1997. Lexical access in aphasic and nonaphasic speakers. Psychol Rev. 104:801–838.

Devlin JT, Russell RP, Davis MH, Price CJ, Wilson J, Moss HE, Matthews PM, Tyler LK. 2000. Susceptibility-induced loss of signal: Comparing PET and fMRI on a semantic task. Neuroimage. 11:589–600.

Dickerson BC, Eichenbaum H. 2010. The episodic memory system: Neurocircuitry and disorders. Neuropsychopharmacology.

Dignam J, Copland D, Rawlings A, O'Brien K, Burfein P, Rodriguez AD. 2016. The relationship between novel word learning and anomia treatment success in adults with chronic aphasia. Neuropsychologia. 81:186–197.

Dignam JK, Rodriguez AD, Copland DA. 2016. Evidence for Intensive Aphasia Therapy: Consideration of Theories From Neuroscience and Cognitive

Psychology. PM R.

Dörfel D, Werner A, Schaefer M, Von Kummer R, Karl A. 2009. Distinct brain networks in recognition memory share a defined region in the precuneus. Eur J Neurosci. 30:1947–1959.

Draganski B, Gaser N, Busch V, Schuierer G, Bogdahn U, May A. 2004. Changes in grey matter induced by training. Nature. 427:311–312.

Dronkers N, Baldo J V. 2009. Language: Aphasia. In: Encyclopedia of Neuroscience. Elsevier. p. 343–348.

Dronkers NF, Plaisant O, Iba-Zizen MT, Cabanis EA. 2007. Paul Broca's historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. Brain. 130:1432–1441.

Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ. 2004. Lesion analysis of the brain areas involved in language comprehension. Cognition. 92:145–177.

Duff MC, Covington N V., Hilverman C, Cohen NJ. 2020. Semantic Memory and the Hippocampus: Revisiting, Reaffirming, and Extending the Reach of Their Critical Relationship. Front Hum Neurosci. 0:471.

Dumay N, Gaskell MG. 2007. Sleep-associated changes in the mental representation of spoken words. Psychol Sci. 18:35–39.

Eichenbaum H, Schoenbaum G, Young B, Bunsey M. 1996. Functional organization of the hippocampal memory system. Proc Natl Acad Sci. 93:13500–13507.

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

Fiez JA. 2001. Neuroimaging studies of speech: An overview of techniques and methodological approaches. J Commun Disord. 34:445–454.

Fillingham JK, Sage K, Lambon Ralph † MA. 2006. The treatment of anomia using errorless learning. Neuropsychol Rehabil. 16:129–154.

Flinker A, Korzeniewska A, Shestyuk AY, Franaszczuk PJ, Dronkers NF, Knight

RT, Crone NE. 2015. Redefining the role of broca's area in speech. Proc Natl Acad Sci U S A. 112:2871–2875.

Forkel SJ, De Schotten MT, Dell'Acqua F, Kalra L, Murphy DGM, Williams SCR, Catani M. 2014. Anatomical predictors of aphasia recovery: a tractography study of bilateral perisylvian language networks. Brain. 137:2027–2039.

Foygel D, Dell GS. 2000. Models of Impaired Lexical Access in Speech Production. J Mem Lang. 43:182–216.

Frankford SA, Nieto-Castañón A, Tourville JA, Guenther FH. 2021. Reliability of single-subject neural activation patterns in speech production tasks. Brain Lang. 212.

Fridriksson J. 2010. Preservation and modulation of specific left hemisphere regions is vital for treated recovery from anomia in stroke. J Neurosci. 30:11558–11564.

Fridriksson J, Moser D, Bonilha L, Morrow-odom KL, Fridriksson A, Baylis GC, Rorden C. 2008. Treatment in Aphasia. 45:1812–1822.

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

Friston KJ, Holmes AP, Worsley KJ, Poline J-P, Frith CD, Frackowiak RSJ. 1994. Statistical parametric maps in functional imaging: A general linear approach. Hum Brain Mapp. 2:189–210.

Gabriel Z, Bowling A. 2004. Quality of life from the perspectives of older people. Ageing Soc. 24:675–691.

Gajardo-Vidal A, Lorca-Puls DL, Hope TMH, Jones OP, Seghier ML, Prejawa S, Crinion JT, Leff AP, Green DW, Price CJ. 2018. How right hemisphere damage after stroke can impair speech comprehension. Brain. 141:3389–3404.

Gaskell MG, Cairney SA, Rodd JM. 2019. Contextual priming of word meanings is stabilized over sleep. Cognition. 182:109–126.

Gaskell MG, Dumay N. 2003. Lexical competition and the acquisition of novel

words. Cognition. 89:105–132.

Geschwind N. 1965. Disconnexion syndromes in animals and man. I. Brain. 88:237–294.

Geva S, Bennett S, Warburton EA, Patterson K. 2011. Discrepancy between inner and overt speech: Implications for post-stroke aphasia and normal language processing. Aphasiology. 25:323–343.

Gilmore AW, Quach A, Kalinowski SE, González-Araya EI, Gotts SJ, Schacter DL, Martin A. 2021. Evidence supporting a time-limited hippocampal role in retrieving autobiographical memories. Proc Natl Acad Sci. 118:2023069118.

Giorgio A, Santelli L, Tomassini V, Bosnell R, Smith S, De Stefano N, Johansen-Berg H. 2010. Age-related changes in grey and white matter structure throughout adulthood. Neuroimage. 51:943–951.

Golestani N, Pallier C. 2007. Anatomical Correlates of Foreign Speech Sound Production. Cereb Cortex. 17:929–934.

Gonzalez-Castillo J, Panwar P, Buchanan LC, Caballero-Gaudes C, Handwerker DA, Jangraw DC, Zachariou V, Inati S, Roopchansingh V, Derbyshire JA, Bandettini PA. 2016. Evaluation of multi-echo ICA denoising for task based fMRI studies: Block designs, rapid event-related designs, and cardiac-gated fMRI. Neuroimage. 141:452.

Good CD, Johnsrude IS, Ashburner J, Henson RNA, Friston KJ, Frackowiak RSJ. 2001. A voxel-based morphometric study of ageing in 465 normal adult human brains. Neuroimage. 14:21–36.

Goodglass H, Kaplan E, Barresi B. 2001. Boston diagnostic aphasia examination. Philadelphia: Lippincott Williams & Wilkins.

Goodglass H, Wingfield A. 1997. Word-finding deficits in aphasia: Brain–behavior relations and clinical symptomatology. In: Goodglass H,, Wingfield A, editors. Anomia: Neuroanatomical and cognitive correlates. San Diego, CA: Academic Press.

Gore KR, Woollams AM, Bruehl S, Halai AD, Lambon Ralph MA. 2021. Direct Neural Evidence for the Contrastive Roles of the Complementary Learning

Systems in Adult Acquisition of Native Vocabulary. Cereb Cortex. 2021:1–14.

Grady CL, Rieck JR, Nichol D, Rodrigue KM, Kennedy KM. 2021. Influence of sample size and analytic approach on stability and interpretation of brain-behavior correlations in task-related fMRI data. Hum Brain Mapp. 42:204–219.

Green DW, Crinion J, Price CJ. 2007. Exploring cross-linguistic vocabulary effects on brain structures using voxel-based morphometry. Biling (Camb Engl). 10:189–199.

Grönholm P, Rinne JO, Vorobyev V, Laine M. 2005. Naming of newly learned objects: A PET activation study. Cogn Brain Res. 25:359–371.

Gupta P, Abbs B, Schwartz M, Lipinski J. 2006. New word learning in aphasic patients: Dissociating phonological and semantic components. Brain Lang. 99:8–9.

Halai AD, Parkes LM, Welbourne SR. 2015. Dual-echo fMRI can detect activations in inferior temporal lobe during intelligible speech comprehension. Neuroimage. 122:214–221.

Halai AD, Welbourne SR, Embleton K, Parkes LM. 2014. A comparison of dual gradient-echo and spin-echo fMRI of the inferior temporal lobe. Hum Brain Mapp. 35:4118–4128.

Halai AD, Woollams AM, Lambon Ralph MA. 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.

Hamamé CM, Alario FX, Llorens A, Liégeois-Chauvel C, Trébuchon-Da Fonseca A. 2014. High frequency gamma activity in the left hippocampus predicts visual object naming performance. Brain Lang. 135:104–114.

Hartman J. 1981. Measurement of early spontaneous recovery from aphasia with stroke. Ann Neurol. 9:89–91.

Heiss W-D, Thiel A. 2006. A proposed regional hierarchy in recovery of post-stroke aphasia. Brain Lang. 98:118–123.

Heiss WD, Kessler J, Thiel A, Ghaemi M, Karbe H. 1999. Differential capacity of left and right hemispheric areas for compensation of poststroke aphasia. Ann Neurol. 45:430–438.

Hickok G, Erhard P, Kassubek J, Helms-Tillery AK, Naeve-Velguth S, Strupp JP, Strick PL, Ugurbil K. 2000. A functional magnetic resonance imaging study of the role of left posterior superior temporal gyrus in speech production: implications for the explanation of conduction aphasia. Neurosci Lett. 287:156–160.

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

Hickok G, Poeppel D. 2007. The cortical organization of speech processing. Nat Rev Neurosci.

Hillis AE. 2002. The handbook of adult language disorders : integrating cognitive neuropsychology, neurology, and rehabilitation. Psychology Press.

Hillis AE, Kleinman JT, Newhart M, Heidler-Gary J, Gottesman R, Barker PB, Aldrich E, Llinas R, Wityk R, Chaudhry P. 2006. Restoring Cerebral Blood Flow Reveals Neural Regions Critical for Naming. J Neurosci. 26:8069–8073.

Hodges JR, Patterson K, Oxbury S, Funnell E. 1992. SEMANTIC DEMENTIA: PROGRESSIVE FLUENT APHASIA WITH TEMPORAL LOBE ATROPHY. Brain. 115:1783–1806.

Hoffman P, McClelland JL, Lambon Ralph MA. 2018. Concepts, control, and context: A connectionist account of normal and disordered semantic cognition. Psychol Rev. 125:293–328.

Hoffman P, Morcom AM. 2018. Age-related changes in the neural networks supporting semantic cognition: A meta-analysis of 47 functional neuroimaging studies. Neurosci Biobehav Rev.

Hoffman P, Pobric G, Drakesmith M, Lambon Ralph MA. 2012. Posterior middle temporal gyrus is involved in verbal and non-verbal semantic cognition: Evidence from rTMS. Aphasiology. 26:1119–1130.

Howard D. 2003. Single Cases, Group Studies and Case Series in Aphasia Therapy. Sci Aphasia. 245–258.

Howard D, Best W, Nickels L. 2015. Optimising the design of intervention studies: critiques and ways forward. Aphasiology. 29:526–562.

Humphreys GF, Jackson RL, Lambon Ralph MA. 2020. Overarching Principles and Dimensions of the Functional Organization in the Inferior Parietal Cortex. Cereb Cortex. 30:5639–5653.

Humphreys GF, Lambon Ralph MA. 2015. Fusion and Fission of Cognitive Functions in the Human Parietal Cortex. Cereb Cortex. 25:3547–3560.

Humphreys GW, Price CJ, Riddoch MJ. 1999. From objects to names: A cognitive neuroscience approach. Psychol Res. 62:118–130.

Indefrey P, Levelt WJ. 2004. The spatial and temporal signatures of word production components. Cognition. 92:101–144.

Ingram RU, Halai AD, Pobric G, Sajjadi S, Patterson K, Lambon Ralph MA. 2020. Graded, multidimensional intra- and intergroup variations in primary progressive aphasia and post-stroke aphasia. Brain. 143:3121–3135.

Jackson RL. 2021. The neural correlates of semantic control revisited. Neuroimage. 224:117444.

Jackson RL, Hoffman P, Pobric G, Lambon Ralph MA. 2016. The semantic network at work and rest: Differential connectivity of anterior temporal lobe subregions. J Neurosci. 36:1490–1501.

Jackson RL, Rogers TT, Lambon Ralph MA. 2021. Reverse-engineering the cortical architecture for controlled semantic cognition. Nat Hum Behav.

Jbabdi S, Sotiropoulos SN, Savio AM, Graña M, Behrens TEJ. 2012. Model-based analysis of multishell diffusion MR data for tractography: how to get over fitting problems. Magn Reson Med. 68:1846–1855.

Jefferies E. 2013. The neural basis of semantic cognition: Converging evidence from neuropsychology, neuroimaging and TMS. Cortex.

Jefferies E, Crisp J, Lambon Ralph MA. 2006. The impact of phonological or

semantic impairment on delayed auditory repetition: Evidence from stroke aphasia and semantic dementia. Aphasiology. 20:963–992.

Jefferies E, Lambon Ralph MA. 2006a. Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. Brain. 129:2132–2147.

Jefferies E, Patterson K, Jones RW, Lambon Ralph MA. 2009. Comprehension of concrete and abstract words in semantic dementia. Neuropsychology. 23:492.

Kaplan E, Goodglass H, Weintraub S. 1983. Boston Naming Test. Philadelphia: Lea & Febiger.

Kay J, Lesser R, Coltheart M. 1992. PALPA: Psycholinguistic assessments of language processing in aphasia. PALPA. 686.

Kumaran D, Hassabis D, Mcclelland JL. 2016. What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. Trends Cogn Sci. 20.

Kundu P, Brenowitz ND, Voon V, Worbe Y, Vertes PE, Inati SJ, Saad ZS, Bandettini PA, Bullmore ET. 2013. Integrated strategy for improving functional connectivity mapping using multiecho fMRI. Proc Natl Acad Sci. 110:16187–16192.

Kundu P, Inati SJ, Evans JW, Luh W-M, Bandettini PA. 2012. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. Neuroimage. 60:1759–1770.

Kundu P, Voon V, Balchandani P, Lombardo M V., Poser BA, Bandettini PA. 2017. Multi-echo fMRI: A review of applications in fMRI denoising and analysis of BOLD signals. Neuroimage. 154:59–80.

Laine M, Salmelin R. 2010. Neurocognition of New Word Learning in the Native Tongue: Lessons From the Ancient Farming Equipment Paradigm. Lang Learn. 60:25–44.

Lambon Ralph MA. 1998. Distributed versus localist representations: Evidence from a study of item consistency in a case of classical anomia. Brain Lang. 64:339–360.

Lambon Ralph MA. 2014. Neurocognitive insights on conceptual knowledge and its breakdown. Philos Trans R Soc B Biol Sci. 369:20120392.

Lambon Ralph MA, Jefferies E, Patterson K, Rogers TT. 2016. The neural and computational bases of semantic cognition. Nat Rev Neurosci.

Lambon Ralph MA, Sage K, Roberts J. 2000. Classical anomia: a neuropsychological perspective on speech production. Neuropsychologia. 38:186–202.

Lambon Ralph MA, Snell C, Fillingham JK, 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. 20:289–305.

Lazar RM, Mohr JP. 2011. Revisiting the contributions of Paul Broca to the study of aphasia. Neuropsychol Rev.

Lazar RM, Speizer AE, Festa JR, Krakauer JW, Marshall RS. 2008. Variability in language recovery after first-time stroke. J Neurol Neurosurg Psychiatry. 79:530–534.

Lee HL, Devlin JT, Shakeshaft C, Stewart LH, Brennan A, Glensman J, Pitcher K, Crinion J, Mechelli A, Frackowiak RSJ, Green DW, Price CJ. 2007. Anatomical Traces of Vocabulary Acquisition in the Adolescent Brain. J Neurosci. 27:1184–1189.

Leff AP, Crinion JT, Holland SK, Turkheimer FE, Howard D, Wise RJS. 2002. A physiological change in the homotopic cortex following left posterior temporal lobe infarction. Ann Neurol. 51:553–558.

Léger A, Démonet JF, Ruff S, Aithamon B, Touyeras B, Puel M, Boulanouar K, Cardebat D. 2002. Neural substrates of spoken language rehabilitation in an aphasic patient: an fMRI study. Neuroimage. 17:174–183.

Lendrem W, Lincoln NB. 1985. Spontaneous recovery of language in patients with aphasia between 4 and 34 weeks after stroke. J Neurol Neurosurg Psychiatry. 48:743–748.

Levelt WJM. 1989. Speaking: From Intention to Articulation. Cambridge: The MIT Press.

Lewis PA, Durrant SJ. 2011. Overlapping memory replay during sleep builds cognitive schemata. Trends Cogn Sci. 15:343–351.

Lichtheim L. 1885. On aphasia. Brain. 7:433–485.

López-Barroso D, Catani M, Ripollés P, Dell'Acqua F, Rodríguez-Fornells A, De Diego-Balaguer R. 2013. Word learning is mediated by the left arcuate fasciculus. Proc Natl Acad Sci U S A. 110:13168–13173.

Lorca-Puls DL, Gajardo-Vidal A, White J, Seghier ML, Leff AP, Green DW, Crinion JT, Ludersdorfer P, Hope TMH, Bowman H, Price CJ. 2018. The impact of sample size on the reproducibility of voxel-based lesion-deficit mappings. Neuropsychologia. 115:101–111.

Lövdén M, Wenger E, Mårtensson J, Lindenberger U, Bäckman L. 2013. Structural brain plasticity in adult learning and development. Neurosci Biobehav Rev. 37:2296–2310.

Manns JR, Hopkins RO, Squire LR. 2003. Semantic memory and the human hippocampus. Neuron. 38:127–133.

Manring ND, Johnson RE. 1996. Modeling and designing a variable-displacement open-loop pump. J Dyn Syst Meas Control Trans ASME. 118:267–271.

Marr D. 1971. Simple memory: a theory for archicortex. Philos Trans R Soc Lond B Biol Sci. 262:23–81.

Martin A, Wiggs CL, Weisberg J. 1997. Modulation of human medial temporal lobe activity by form, meaning, and experience. Hippocampus. 7:587–593.

McClelland JL. 1979. On the time relations of mental processes: An examination of systems of processes in cascade. Psychol Rev. 86:287–330.

McClelland JL. 2013. Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. J Exp Psychol Gen. 142:1190–1210.

McClelland JL, McNaughton BL, Lampinen AK. 2020. Integration of new information in memory: new insights from a complementary learning systems perspective. Philos Trans R Soc B Biol Sci. 375:20190637.

McClelland JL, McNaughton BL, O'Reilly RC. 1995. Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychologi. 102:419–457.

McClelland JL, Rumelhart DE, Group PR. 1986. Parallel Distributed Processing. Cambridge, MA: MIT Press.

McCloskey M, Cohen NJ. 1989. Catastrophic Interference in Connectionist Networks: The Sequential Learning Problem. Psychol Learn Motiv - Adv Res Theory. 24:109–165.

Mechelli A, Crinion JT, Noppeney U, O' Doherty J, Ashburner J, Frackowiak RS, Price CJ. 2004. Structural plasticity in the bilingual brain. Nat 2004 4317010. 431:757–757.

Meinzer M, Breitenstein C. 2008. Functional imaging studies of treatment-induced recovery in chronic aphasia. Aphasiology. 22:1251–1268.

Mestres-Missé A, Càmara E, Rodriguez-Fornells A, Rotte M, Münte TF. 2008. Functional neuroanatomy of meaning acquisition from context. J Cogn Neurosci. 20:2153–2166.

Mestres-Missé A, Rodriguez-Fornells A, Münte TF. 2007. Watching the brain during meaning acquisition. Cereb Cortex. 17:1858–1866.

Mesulam MM, Thompson CK, Weintraub S, Rogalski EJ. 2015. The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. Brain. 138:2423–2437.

Minkina I, Rosenberg S, Kalinyak-Fliszar M, Martin N. 2017. Short-Term Memory and Aphasia: From Theory to Treatment. Semin Speech Lang. 38:17.

Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR. 2006. The Addenbrooke's Cognitive Examination revised (ACE-R): A brief cognitive test battery for dementia screening. Int J Geriatr Psychiatry. 21:1078–1085.

Moore CJ, Price CJ. 1999. Three distinct ventral occipitotemporal regions for reading and object naming. Neuroimage. 10:181–192.

Murtha S, Chertkow H, Beauregard M, Evans A. 1999. The Neural Substrate of Picture Naming. J Cogn Neurosci. 11:399–423.

Nadel L, Moscovitch M. 1997. Memory consolidation, retrograde amnesia and the hippocampal complex. Curr Opin Neurobiol. 7:217–227.

Naeser MA, Martin PI, Baker EH, Hodge SM, Sczerzenie SE, Nicholas M, Palumbo CL, Goodglass H, Wingfield A, Samaraweera R, Harris G, Baird A, Renshaw P, Yurgelun-Todd D. 2004. Overt propositional speech in chronic nonfluent aphasia studied with the dynamic susceptibility contrast fMRI method. Neuroimage. 22:29–41.

Naeser MA, Martin PI, Nicholas M, Baker EH, Seekins H, Helm-Estabrooks N, Cayer-Meade C, Kobayashi M, Theoret H, Fregni F, Tormos JM, Kurland J, Doron KW, Pascual-Leone A. 2005. Improved naming after TMS treatments in a chronic, global aphasia patient--case report. Neurocase. 11:182–193.

Nardo D, Holland R, Leff AP, Price CJ, Crinion JT. 2017. Less is more: neural mechanisms underlying anomia treatment in chronic aphasic patients.

Nickels L. 2002. Therapy for naming disorders: Revisiting, revising, and reviewing. Aphasiology. 16:935–979.

Noonan KA, Jefferies E, Corbett F, Lambon Ralph MA. 2010. Elucidating the Nature of Deregulated Semantic Cognition in Semantic Aphasia: Evidence for the Roles of Prefrontal and Temporo-parietal Cortices. J Cogn Neurosci. 22:1597–1613.

Noonan KA, Jefferies E, Visser M, Lambon Ralph MA. 2013. Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. J Cogn Neurosci. 25:1824–1850.

Norman KA, O'reilly RC. 2003. Modeling Hippocampal and Neocortical Contributions to Recognition Memory: A Complementary-Learning-Systems Approach.

O'Reilly JX, Woolrich MW, Behrens TEJ, Smith SM, Johansen-Berg H. 2012. Tools of the trade: psychophysiological interactions and functional connectivity. Soc

Cogn Affect Neurosci. 7:604.

Ozubko JD, Joordens S. 2011. The Similarities (and Familiarities) of Pseudowords and Extremely High-Frequency Words: Examining a Familiarity-Based Explanation of the Pseudoword Effect. J Exp Psychol Learn Mem Cogn. 37:123–139.

Palmer R, Dimairo M, Cooper C, Enderby P, Brady M, Bowen A, Latimer N, Julious S, Cross E, Alshreef A, Harrison M, Bradley E, Witts H, Chater T. 2019. Self-managed, computerised speech and language therapy for patients with chronic aphasia post-stroke compared with usual care or attention control (Big CACTUS): a multicentre, single-blinded, randomised controlled trial. Lancet Neurol. 18:821–833.

Palmer R, Hughes H, Chater T. 2017. What do people with aphasia want to be able to say? A content analysis of words identified as personally relevant by people with aphasia. PLoS One. 12:e0174065.

Patenaude B, Smith SM, Kennedy DN, Jenkinson M. 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage. 56:907–922.

Patterson K, Nestor PJ, Rogers TT. 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. Nat Rev Neurosci.

Paulesu E, Vallar G, Berlingeri M, Signorini M, Vitali P, Burani C, Perani D, Fazio F. 2009. Supercalifragilisticexpialidocious: How the brain learns words never heard before. Neuroimage. 45:1368–1377.

Peñaloza C, Mirman D, Tuomiranta L, Benetello A, Heikius IM, Järvinen S, Majos MC, Cardona P, Juncadella M, Laine M, Martin N, Rodríguez-Fornells A. 2016. Novel word acquisition in aphasia: Facing the word-referent ambiguity of natural language learning contexts. Cortex. 79:14–31.

Penny W, Friston K, Ashburner J, Kiebel S, Nichols T. 2007. Statistical Parametric Mapping: The Analysis of Functional Brain Images. Stat Parametr Mapp Anal Funct Brain Images.

Pobric G, Jefferies E, Lambon Ralph MA. 2007. Anterior temporal lobes mediate

semantic representation: Mimicking semantic dementia by using rTMS in normal participants. Proc Natl Acad Sci U S A. 104:20137–20141.

Pobric G, Jefferies E, Lambon Ralph MA. 2010. Amodal semantic representations depend on both anterior temporal lobes: Evidence from repetitive transcranial magnetic stimulation. Neuropsychologia. 48:1336–1342.

Poeppel D. 2014. The neuroanatomic and neurophysiological infrastructure for speech and language. Curr Opin Neurobiol. 28:142–149.

Pohl A, Meyer CM, Deckers A, Eckmann T, Barthel J, Huber W, Habel U, Willmes K, Abel S. 2017. Training-related changes of brain activation for speech production in healthy speakers–a longitudinal fMRI study to mimic aphasia therapy. Aphasiology. 31:495–518.

Poser BA, Versluis MJ, Hoogduin JM, Norris DG. 2006. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. Magn Reson Med. 55:1227–1235.

Posse S, Wiese S, Gembris D, Mathiak K, Kessler C, Grosse-Ruyken ML, Elghahwagi B, Richards T, Dager SR, Kiselev VG. 1999. Enhancement of BOLD-contrast sensitivity by single-shot multi-echo functional MR imaging. Magn Reson Med. 42:87–97.

Price C. 2011. A generative model of speech production in Broca's and Wernicke's areas. Front Psychol. 2:237.

Price C, Thierry G, Griffiths T. 2005. Speech-specific auditory processing: where is it? Trends Cogn Sci. 9:271–276.

Price CJ. 2010. The anatomy of language: a review of 100 fMRI studies published in 2009. Ann N Y Acad Sci. 1191:62–88.

Price CJ. 2012. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. Neuroimage. 62:816–847.

Price CJ, Devlin JT. 2003. The myth of the visual word form area. Neuroimage. 19:473–481.

Price CJ, Moore CJ, Humphreys GW, Frackowiak RSJ, Friston KJ. 1996. The neural

regions sustaining object recognition and naming. Proc R Soc B Biol Sci. 263:1501–1507.

Pustina D, Coslett HB, Turkeltaub PE, Tustison N, Schwartz MF, Avants B. 2016. Automated segmentation of chronic stroke lesions using LINDA: Lesion identification with neighborhood data analysis. Hum Brain Mapp. 37:1405.

Raboyeau G, Marie N, Balduyck S, Gros H, Démonet JF, Cardebat D. 2004. Lexical learning of the English language: A PET study in healthy French subjects. Neuroimage. 22:1808–1818.

Raven J. 1962. Coloured progressive matrices. Sets A, Ab & B. London: H. K. Lewis.

Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, Dahle C, Gerstorf D, Acker JD. 2005. Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. Cereb Cortex. 15:1676–1689.

Resnick SM, Pham DL, Kraut MA, Zonderman AB, Davatzikos C. 2003. Longitudinal magnetic resonance imaging studies of older adults: A shrinking brain. J Neurosci. 23:3295–3301.

Ribot T. 1881. Les maladies de la memoire. Paris: Germer Baillare.

Ribot T. 1882. Diseases of memory. New York, NY: D. Appleton and Company.

Rice GE, Caswell H, Moore P, Hoffman P, Lambon Ralph MA. 2018. The roles of left versus right anterior temporal lobes in semantic memory: A neuropsychological comparison of postsurgical temporal lobe epilepsy patients. Cereb Cortex. 28:1487–1501.

Rodd JM, Davis MH, Johnsrude IS. 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb Cortex. 15:1261–1269.

Rodriguez-Fornells A, Cunillera T, Mestres-Missé A, De Diego-Balaguer R. 2009. Neurophysiological mechanisms involved in language learning in adults. Philos Trans R Soc B Biol Sci. 364:3711.

Rogers TT, Lambon Ralph MA, Garrard P, Bozeat S, McClelland JL, Hodges JR, Patterson K. 2004. Structure and Deterioration of Semantic Memory: A Neuropsychological and Computational Investigation. Psychol Rev. 111:205–235.

Rohrer JD, Knight WD, Warren JE, Fox NC, Rossor MN, Warren JD. 2008. Word-finding difficulty: A clinical analysis of the progressive aphasias. Brain. 131:8–38.

Rolls ET, Huang CC, Lin CP, Feng J, Joliot M. 2020. Automated anatomical labelling atlas 3. Neuroimage. 206:116189.

Rosen HJ, Petersen SE, Linenweber MR, Snyder AZ, White DA, Chapman L, Dromerick AW, Fiez JA, Corbetta, MD. 2000. Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. Neurology. 55:1883–1894.

Rosenbaum RS, Moscovitch M, Foster JK, Schnyer DM, Gao F, Kovacevic N, Verfaellie M, Black S, Levine B. 2008. Patterns of autobiographical memory loss in medial-temporal lobe amnesic patients. J Cogn Neurosci. 20:1490–1506.

Ross ED. 1985. The Assessment of Aphasia and Related Disorders, Neurology. Philadelphia: Lea & Febiger.

Ross Graham J, Pereira S, Teasell R. 2011. Aphasia and return to work in younger stroke survivors. Aphasiology. 25:952–960.

RStudioTeam. 2020. Integrated development for R. RStudio.

Rumelhart DE, Hinton GE, Williams RJ. 1986. Learning representations by back-propagating errors. Nat 1986 3236088. 323:533–536.

Sanjuán A, Hope TMH, Parker Jones ’Oiwi, Prejawa S, Oberhuber M, Guerin J, Seghier ML, Green DW, Price CJ. 2015. Dissociating the semantic function of two neighbouring subregions in the left lateral anterior temporal lobe. Neuropsychologia. 76:153–162.

Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Vry M-S, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C. 2008. Ventral and dorsal pathways for language. Proc Natl Acad Sci U S A.

105:18035–18040.

Saur D, Lange R, Baumgaertner A, Schraknepper V, Willmes K, Rijntjes M, Weiller C. 2006. Dynamics of language reorganization after stroke. Brain. 129:1371–1384.

Schapiro AC, McClelland JL, Welbourne SR, Rogers TT, Ralph MAL. 2013. Why Bilateral Damage Is Worse than Unilateral Damage to the Brain. J Cogn Neurosci. 25:2107–2123.

Schapiro AC, McDevitt EA, Rogers TT, Mednick SC, Norman KA. 2018. Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance. Nat Commun 2018 91. 9:1–11.

Schapiro AC, Turk-Browne NB, Botvinick MM, Norman KA. 2017. Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. Philos Trans R Soc B Biol Sci. 372.

Schwartz M, Dell G, Martin N, Gahl S, Sobel P. 2006. A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. J Mem Lang. 54:228–264.

Schwartz MF, Dell GS. 2012. Case series investigations. Congitive Neuropsychol. 27:477–494.

Scoville WB, Milner B. 1957. Loss of recent memory after bilateral hippocampal lesions. J Neurol Neurosurg Psychiatry. 20:11–21.

Seghier ML, Ramlackhansingh A, Crinion J, Leff AP, Price CJ. 2008. Lesion identification using unified segmentation-normalisation models and fuzzy clustering. Neuroimage. 41:1253–1266.

Seidenberg MS, McClelland JL. 1989. A Distributed, Developmental Model of Word Recognition and Naming. Psychol Rev. 96:523–568.

Sekeres MJ, Winocur G, Moscovitch M. 2018. The hippocampus and related neocortical structures in memory transformation. Neurosci Lett. 680:39–53.

Shallice T. 1988. From neuropsychology to mental structure. Cambridge, UK:

Cambridge University Press.

Shimotake A, Matsumoto R, Ueno T, Kunieda T, Saito S, Hoffman P, Kikuchi T, Fukuyama H, Miyamoto S, Takahashi R, Ikeda A, Lambon Ralph MA. 2015. Direct exploration of the role of the ventral anterior temporal lobe in semantic memory: Cortical stimulation and local field potential evidence from subdural grid electrodes. Cereb Cortex. 25:3802–3817.

Sliwinska MW, Khadilkar M, Campbell-Ratcliffe J, Quevenco F, Devlin JT. 2012. Early and Sustained Supramarginal Gyrus Contributions to Phonological Processing. Front Psychol. 3.

Smith SM. 2002. Fast robust automated brain extraction. Hum Brain Mapp. 17:143–155.

Solomon EA, Lega BC, Sperling MR, Kahana MJ. 2019. Hippocampal theta codes for distances in semantic and temporal spaces. Proc Natl Acad Sci U S A. 116:24343–24352.

Sowell ER, Peterson BS, Thompson PM, Welcome SE, Henkenius AL, Toga AW. 2003. Mapping cortical change across the human life span. Nat Neurosci. 6:309–315.

Squire LR, Alvarez P. 1995. Retrograde amnesia and memory consolidation: a neurobiological perspective. Curr Opin Neurobiol. 5:169–177.

Squire LR, Stark CEL, Clark RE. 2004. The Medial Temporal Lobe. Annu Rev Neurosci. 27:279–306.

St. John MF, McClelland JL. 1990. Learning and applying contextual constraints in sentence comprehension. Artif Intell. 46:217–257.

Stefaniak JD, Alyahya RSW, Lambon Ralph MA. 2021. Language networks in aphasia and health: A 1000 participant activation likelihood estimation meta-analysis. Neuroimage. 233:117960.

Stefaniak JD, Halai AD, Lambon Ralph MA. 2020. The neural and neurocomputational bases of recovery from post-stroke aphasia. Nat Rev Neurol.

Swinburn K, Porter G, Howard D. 2005. Comprehensive aphasia test. Hove: Psychology Press.

Szaflarski JP, Allendorfer JB, Banks C, Vannest J, Holland SK. 2013. Recovered vs. not-recovered from post-stroke aphasia: The contributions from the dominant and non-dominant hemispheres. Restor Neurol Neurosci. 31:347–360.

Takashima A, Bakker I, van Hell J, Janzen G, McQueen J. 2014. Richness of information about novel words influences how episodic and semantic memory networks interact during lexicalization. Neuroimage. 84:265–278.

Takashima A, Bakker I, van Hell JG, Janzen G, McQueen JM. 2017. Interaction between episodic and semantic memory networks in the acquisition and consolidation of novel spoken words. Brain Lang. 167:44–60.

Tamminen J, Lambon Ralph MA, Lewis PA. 2013. The Role of Sleep Spindles and Slow-Wave Activity in Integrating New Information in Semantic Memory. J Neurosci. 33:15376–15381.

Tamminen J, Payne JD, Stickgold R, Wamsley EJ, Gaskell MG. 2010. Sleep spindle activity is associated with the integration of new memories and existing knowledge. J Neurosci. 30:14356–14360.

Taylor JSH, Davis MH, Rastle K. 2019. Mapping visual symbols onto spoken language along the ventral visual stream. Proc Natl Acad Sci U S A. 116:17723–17728.

Taylor JSH, Rastle K, Davis MH. 2013. Can cognitive models explain brain activation during word and pseudoword reading? A meta-analysis of 36 neuroimaging studies. Psychol Bull. 139:766–791.

Taylor JSH, Rastle K, Davis MH. 2014. Interpreting response time effects in functional imaging studies. Neuroimage. 99:419.

Thiel A, Habedank B, Herholz K, Kessler J, Winhuisen L, Haupt WF, Heiss WD. 2006. From the left to the right: How the brain compensates progressive loss of language function. Brain Lang. 98:57–65.

Thompson K, Gibbings A, Shaw J, Ray L, Hébert G, De Koninck J, Fogel S. 2021. Sleep and Second-Language Acquisition Revisited: The Role of Sleep Spindles

and Rapid Eye Movements. Nat Sci Sleep. 13:1887.

Townsend J, Ashby F. 1978. Methods of modeling capacity in simple processing systems. In: Castellan J,, Restle F, editors. Cognitive Theory. Vol. 3. Erlbaum: Hillsdale, NJ. p. 200–239.

Tulving E. 1972. Episodic and semantic memory. In: Organization of Memory. Academic Press.

Tulving E. 1985. How many memory systems are there? Am Psychol. 40:385–398.

Tulving E, Markowitsch HJ. 1998. Episodic and declarative memory: Role of the hippocampus. Hippocampus.

Tuomiranta L, Grönroos AM, Martin N, Laine M. 2014. Vocabulary acquisition in aphasia: Modality can matter. J Neurolinguistics. 32:42–58.

Tuomiranta L, Rautakoski P, Rinne JO, Martin N, Laine M. 2012. Long-term maintenance of novel vocabulary in persons with chronic aphasia. Aphasiology. 26:1053–1073.

Tuomiranta LM, Càmara E, Froudist Walsh S, Ripollés P, Saunavaara JP, Parkkola R, Martin N, Rodríguez-Fornells A, Laine M. 2015. Hidden word learning capacity through orthography in aphasia. Cortex. 50:174–191.

Turkeltaub PE, Coslett HB, Thomas AL, Faseyitan O, Benson J, C N, Hamilton RH. 2012. The right hemisphere is not unitary in its role in aphasia recovery. Cortex. 48:1179–1186.

Turken AU, Dronkers NF. 2011. The Neural Architecture of the Language Comprehension Network: Converging Evidence from Lesion and Connectivity Analyses. Front Syst Neurosci. 5.

Tyler LK, Marslen-Wilson W, Stamatakis EA. 2005. Dissociating neuro-cognitive component processes: Voxel-based correlational methodology. Neuropsychologia. 43:771–778.

Tyler LK, Shafto MA, Randall B, Wright P, Marslen-Wilson WD, Stamatakis EA. 2010. Preserving syntactic processing across the adult life span: The modulation of the frontotemporal language system in the context of age-related atrophy.

Cereb Cortex. 20:352–364.

Ueno T, Saito S, Rogers TT, Lambon Ralph MA. 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 Hees S, McMahon K, Angwin A, de Zubicaray G, Copland DA. 2014. Neural activity associated with semantic versus phonological anomia treatments in aphasia. Brain Lang. 129:47–57.

van Oers C, Vink M, van Zandvoort M, van der Worp H, de Haan E, Kappelle L, Ramsey N, Dijkhuizen R. 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. 49:885–893.

Vilberg KL, Rugg MD. 2008. Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective. Neuropsychologia. 46:1787–1799.

Visser M, Jefferies E, Embleton K V., Lambon Ralph MA. 2012. Both the middle temporal gyrus and the ventral anterior temporal area are crucial for multimodal semantic processing: Distortion-corrected fMRI evidence for a double gradient of information convergence in the temporal lobes. J Cogn Neurosci. 24:1766–1778.

Visser M, Jefferies E, Lambon Ralph MA. 2010. Semantic Processing in the Anterior Temporal Lobes: A Meta-analysis of the Functional Neuroimaging Literature. J Cogn Neurosci. 22:1083–1094.

Wagner AD, Shannon BJ, Kahn I, Buckner RL. 2005. Parietal lobe contributions to episodic memory retrieval. Trends Cogn Sci. 9:445–453.

Warrington EK. 1975b. The Selective Impairment of Semantic Memory. Q J Exp Psychol. 27:635–657.

Wechsler D. 1955. Manual for the Wechsler memory scale-revised. San Antonio: Psychological Corporation.

Welbourne SR, Lambon Ralph MA. 2005. Exploring the impact of plasticity-related recovery after brain damage in a connectionist model of single-word reading.

Cogn Affect Behav Neurosci. 5:77–92.

Welbourne SR, Lambon Ralph M a. 2007. Using parallel distributed processing models to simulate phonological dyslexia: the key role of plasticity-related recovery. J Cogn Neurosci. 19:1125–1139.

Wenger E, Kühn S, Verrel J, Mårtensson J, Bodammer NC, Lindenberger U, Lövdén M. 2017. Repeated Structural Imaging Reveals Nonlinear Progression of Experience-Dependent Volume Changes in Human Motor Cortex. Cereb Cortex. 27:2911–2925.

Wernicke C. 1874. The aphasic symptom-complex. A psychological study on an anatomical basis., Wernicke's works on aphasia. The Hague: Mouton.

Whitney C, Kirk M, O'Sullivan J, Lambon Ralph MA, Jefferies E. 2012. Executive semantic processing is underpinned by a large-scale neural network: revealing the contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. J Cogn Neurosci. 24:133–147.

Wilson MA, McNaughton BL. 1994. Reactivation of Hippocampal Ensemble Memories During Sleep. Science (80-). 265:676–679.

Winhuisen L, Thiel A, Schumacher B, Kessler J, Rudolf J, Haupt WF, Heiss WD. 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. 36:1759–1763.

Wise RJS, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA. 2001. Separate neural subsystems within "Wernicke"s area'. Brain. 124:83–95.

Woodhead ZVJ, Crinion J, Teki S, Penny W, Price CJ, Leff AP. 2017. Auditory training changes temporal lobe connectivity in "Wernicke"s aphasia': A randomised trial. J Neurol Neurosurg Psychiatry. 88:586–594.

Woolrich MW, Jbabdi S, Patenaude B, Chappell M, Makni S, Behrens T, Beckmann C, Jenkinson M, Smith SM. 2009. Bayesian analysis of neuroimaging data in FSL. Neuroimage. 45:S173-86.

Yang J, Gates KM, Molenaar P, Li P. 2015. Neural changes underlying successful second language word learning: An fMRI study. J Neurolinguistics. 33:29–49.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated synthesis of human functional neuroimaging data. Nat Methods. 8:665–670.

Zhang Y, Brady M, Smith S. 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging. 20:45–57.

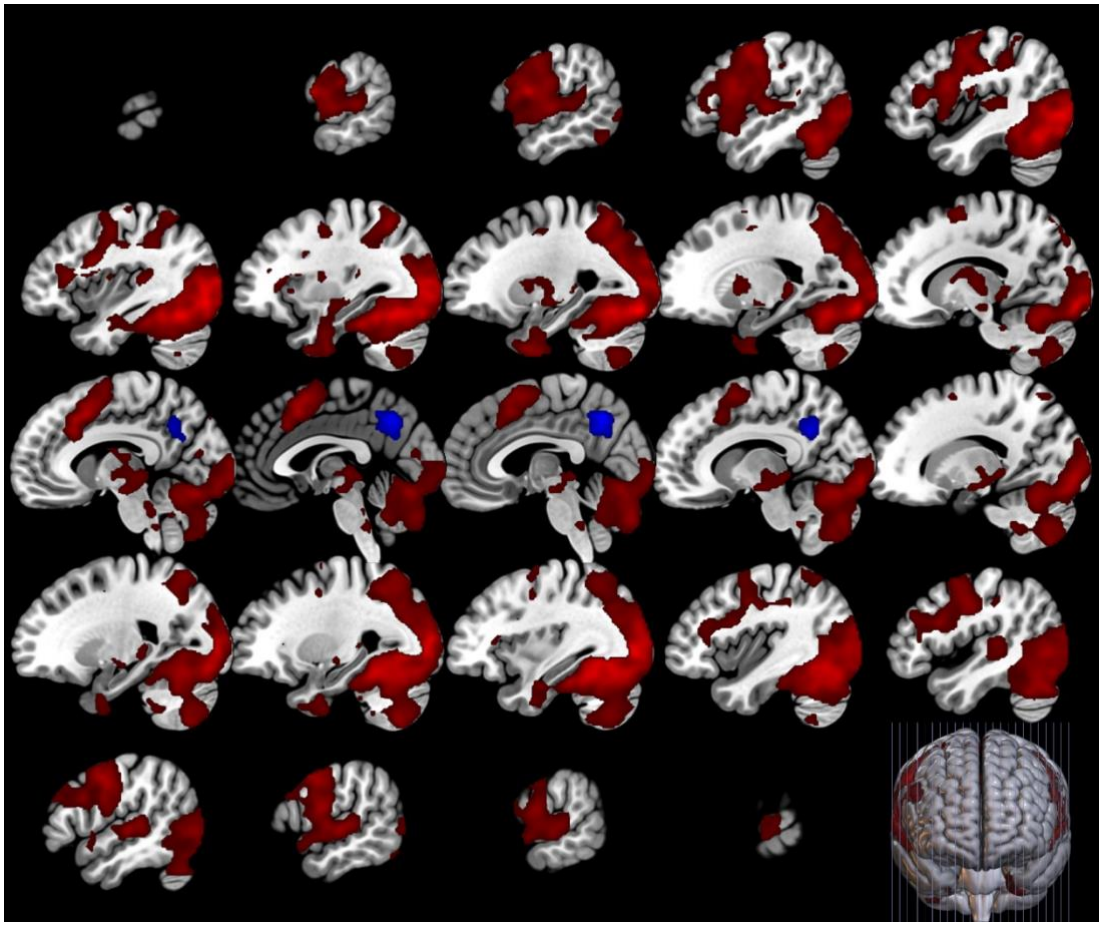
SUPPLEMENTARY DATA: CHAPTER 1

S1. All orthographic stimuli in the study, with 50 known, trained and untrained word sets. Trained and untrained sets were counterbalanced across participants.

Known	Trained	Untrained
Accordion	Aggry	Agouti
Backpack	Ale-warmer	Amphiuma
Bagpipes	Almucantar	Aquamanile
Banjo	Amice	Bismuth
Cactus	Ankus	Blenny
Cauliflower	Anole	Braies
Cheetah	Armet	Cabasset
Cockerel	Aspergillum	Caracal
Codpiece	Astrarium	Carrasow
Crayon	Axolotl	Chapman
Dragonfly	Babirusa	Chevrotain
Dungarees	Bandura	Clepsydra
Earplugs	Bilby	Crinoline
Funnel	Binnacle	Dik-dik
Giraffe	Buckler	Douroucouli
Hairdryer	Cacomistle	Dugong
Hammock	Canezou	Electrophore
Hippopotamus	Celesta	Fipple
Kangaroo	Cervelat	Frogfish
Ladle	Chronometer	Galago
Lawnmower	Civet	Guereza
Mittens	Colugo	Habergeon
Peanut	Corslet	Hawkfish
Pickaxe	Cuscus	Hellbender
Piglet	Dasher	Hirola
Pitchfork	Dilruba	Kakapo
Pliers	Douc	Markhor
Protractor	Echidna	Mattock
Pumpkin	Galligaskins	Mayuri
Radish	Gelada	Micrometer
Saxophone	Gerenuk	Narwhal
Screwdriver	Gharial	Nudibranch
Seahorse	Gorget	Numbat
Seesaw	Headstall	Olm
Sharpener	Hoatzin	Peccary
Shoehorn	Hyrax	Pelerine
Shuttlecock	Kinkajou	Pichiciego
Snowman	Lemming	Polonaise
Spatula	Olinguito	Pooter
Stapler	Paca	Saolo
Starfish	Pangolin	Sengis

Tambourine	Retractors	Shako
Tricycle	Rongeur	Stonefish
Turnip	Saiga	Twybil
Tutu	Solenodon	Vaquita
Tweezers	Tarsier	Versorium
Watermelon	Trepan	Vichitra
Wrench	Turnshoe	Wimple
Xylophone	Ulu	Zibellino
Yo-yo	Vibraphone	Zurna

S2. Significant clusters of activation (red) and deactivation (blue) in the known > rest picture naming contrast. P < .001 voxel level, p < .05 FWE cluster correction.



S3. Significant clusters of activation (red) in the trained > rest picture naming contrast. P < .001 voxel level, p < .05 FWE cluster correction. No significantly deactivated clusters.