Understanding the behavioural and neurological response to digital therapy in chronic post-stroke aphasia

Emily Rachel Charlotte Upton

2022

UCL Queen Square Institute of Neurology

University College London

This thesis is submitted for the degree of

Doctor of Philosophy

Declaration of authorship

I, Emily Upton, confirm that the work presented in my thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis investigates the efficacy of a novel, self-led, tablet-based, word retrieval therapy for persons with chronic post-stroke aphasia (PWA).

In Chapter 1 of this thesis, I analysed data from a Phase II clinical trial of the word retrieval therapy (iTalkBetter), which compared six weeks of therapy with six weeks of standard care. The results showed: (i) PWA made large and significant gains in the retrieval of trained words; (ii) improvements were found for both trained concrete and trained abstract items; (iii) PWA generalised their single word learning to a less constrained task (Spoken Picture Description); (iiii) a combination of baseline factors explained a large amount of the variability in therapy response.

In Chapter 2, I used longitudinal voxel-based morphometry (VBM) to explore therapy-induced structural brain changes in a subset of participants (n=17). I found that taking part in the iTalkBetter therapy programme was associated with increased volume in the grey and white matter of the left and right hemispheres, in the language and cognitive networks.

In Chapter 3, I investigated therapy-driven changes in task-related functional activity in the same subgroup of participants. This analysis showed that changes in activation at post-therapy were correlated with dose of treatment (how much therapy participants completed). These activity changes were found bilaterally in areas supporting auditory analysis and speech production.

In the general discussion section, I discuss the main behavioural, structural and functional findings, limitations, and possible avenues for future research.

Impact statement

There are currently an estimated 1.2 million stroke survivors in the UK and this number is expected to rise by 27% by the year 2047 (Stroke Statistics, 2017; Wafa et al, 2020). One third of stroke survivors will have aphasia, a generalised language disorder which often persists into the chronic stage of recovery (more than six months post-stoke) (Feigin, Norrving & Mensah, 2017; Johansson, Carlsson & Sonnander, 2012). Persons with aphasia (PWA) can have deficits across any and all language modalities and this can have a considerable impact on the ability to communicate, subsequently affecting one's quality of life (Papathanisou, Coppens & Davidson, 2017; Vickers, 2010; Lam & Wodchris, 2010). The research evidence to date indicates that treatments for aphasia are effective and can improve language functioning many years after a stroke occurs, but a large dose of therapy is required (20-100 hours) (Brady et al, 2022; Bhogal, Teasell & Speechley, 2003). Due to the limited rehabilitation resources within the National Health Service, however, PWA often receive just 12 hours of therapy (Clarke et al, 2018; Palmer, Witts & Chater, 2018). Digitising evidence-based therapies offers a cost-effective and feasible approach to provide individuals with the high dosage required to recover language functions.

The first part of my thesis investigated the efficacy of one such digital therapy, 'iTalkBetter', in improving the most common symptom of aphasia, word retrieval deficits. The results from this Phase II clinical trial showed that participants with chronic aphasia made large and significant improvements in the retrieval of trained items following six weeks of therapy. Due to these results, iTalkBetter will now be made available to the public in a Phase III roll-out trial of the therapy. As such, the

testing of this treatment app provides a clear impact to the wider healthcare and PWA community by:

(1) Giving the public access to an evidence-based app for chronic word retrieval deficits.
(2) Enabling PWA to achieve the high doses of therapy required to improve language functions via a self-led, digital intervention.
(3) Providing clinicians with an intervention tool which is feasible to use within the current limited rehabilitation services for post-stroke aphasia.

The findings in my thesis also provide a novel contribution to the study of aphasia therapy by showing structural adaption in the language and cognitive networks following six weeks of iTalkBetter. The longitudinal method used in this study (Ashburner, 2013) has been previously employed to explore structural neuroplasticity induced by a digital intervention for speech comprehension in PWA (Fleming et al, 2020), but this study is the first to do so following word finding therapy. Furthermore, therapy-driven changes in task-related functional activity were observed bilaterally in the language network in relation to how much therapy participants completed; a finding which has not been previously reported in the aphasia literature. These results identify new avenues for future research in aphasia rehabilitation.

Acknowledgements

First and foremost, my thanks goes to Professor Alex Leff for giving me the opportunity to complete a PhD, offering continuous guidance and support, and giving your time and resources to provide generous advice and feedback.

I would also like thank my secondary supervisor, Dr Tom Hope, for your expert recommendations in the structural imaging analyses and interpretations, and for your overall calm and reassuring sentiments that it “will be done”.

I am also grateful to everyone in the Neurotherapeutics team. Thank you for your friendships, laughs, and many trips to the Queen’s Larder over the last three and a half years – it definitely made it all worth it! In particular, thank you to Dr Victoria Fleming for the endless help you have given me throughout this study, and to (future Dr) Tom Langford, my PhD buddy, for keeping me going through the ups and the downs.

A huge thanks is also extended to my family and to Luke, for their unwavering patience and encouragement throughout my academic career. I could not have completed this PhD without your support.

Last, but definitely not least, thank you to all of the participants who to took part in the iTalkBetter trial. Your generosity with your time and your willingness to take part in research made this study possible.

Statement of contribution

The iTalkBetter project was funded by the National Institute for Health Research, via Professor Alex Leff's research professorship. iTalkBetter was developed by Soft-V, an apps development company, alongside Professor Alex Leff, Dr Catherine Doogan, Henry Coley-Fisher, myself, and people with post-stroke aphasia and their carers. Soft-V also programmed the novel word retrieval test following an online experiment I completed with healthy control participants, and I collaborated with a graphic designer to develop the novel spoken picture description scenes. The structural and functional imaging was funded by the Predicting Language Outcome and Recovery After Stroke (PLORAS) research project. The recruitment of participants and the data collection was primarily carried out by myself, with the help of Dr Victoria Fleming and Henry Coley-Fisher. All of the data pre-processing, statistical analyses, and interpretations are my own, and any other contributions have been referenced.

Contents

List of Figures

List of Tables

List of Abbreviations

4-WW	4-Way Weigl
ALI	Automated lesion identification
ALM	Automatic linear modelling
ANOVA	Analysis of variance
aPFC	Anterior prefrontal cortex
ATL	Anterior temporal lobe
BOLD	Blood-Oxygen Level Dependent
CAT	Comprehensive Aphasia Test
CFIT	Cattell Culture Fair Intelligence Test
CSLT	Computerised self-led therapy
dACC	Dorsal anterior cingulate cortex
DKI	Diffusional kurtosis imaging
DLPFC	Dorsolateral prefrontal cortex
DTI	Diffusion tensor imaging
DVAMS	Dynamic Visual Analogue Mood Scales
EPI	Echo planar imaging
FA	Fractional anisotropy
fMRI	Functional magnetic resonance imaging
FWE	Family Wise Error
FWHM	Full-width half maximum
GLM	General linear model
GM	Grey matter
ICW	Information carry words
IFG	Inferior frontal gyrus
ITG	Inferior temporal gyrus
LH	Left hemisphere
M	Mean
MIT	Melodic Intonation Therapy
MNI	Montreal Neurological Institute

MRI	Magnetic resonance imaging
MTG	Middle temporal gyrus
n	Number of participants
NUVA	Naming utterance verification system
PAC	Primary auditory cortex
PALPA	Psycholinguistic Assessments of Language Processing in Aphasia
PCA	Phonological Component Analysis
PLORAS	Predicting Language Outcome and Recovery After Stroke
pSTG	Posterior superior temporal gyrus
PT	Planum temporale
PWA	Persons with aphasia
RH	Right hemisphere
SART	Sustained Attention to Response Task
SD	Standard deviation
SFA	Semantic Feature Analysis
SPD	Spoken Picture Description
SPM12	Statistical parametric mapping software 12
SPSS	Statistical software package for the social sciences
Spt	Sylvian parietal-temporal junction
STG	Superior temporal gyrus
STS	Superior temporal sulcus
tDCS	Transcranial direct current stimulation
TSS	Time since stroke
VBM	Voxel-based morphometry
WM	White matter
WRT	Word Retrieval Test

1 Introduction

The incidence of stroke continues to rise with current estimates around 100,000 every year in the UK (Stroke Statistics, 2017). Medical advances in the acute treatment of stroke have increased the rate of survival, however; as the leading cause of disability, there is an ever-rising number of people living with long-standing impairments. Aphasia, an acquired language disorder, is the second most common major impairment following stroke, affecting around one third of stroke survivors, and can affect any and all modalities of language, including production, comprehension, reading and writing (Feigin, Norrving & Mensah, 2017*;* Papathanisou, Coppens & Davidson, 2017).

As the capacity to communicate effectively is highly integral to the ability to maintain and engage in relationships, aphasia can have devastating consequences. Breakdowns in social networks and relationships can lead to social isolation and poor quality of life, with one study finding that having aphasia had a larger negative impact on quality of life than conditions such as cancer and Alzheimer's disease (Vickers, 2010; Lam & Wodchris, 2010). These implications contribute to a higher incidence of depression and anxiety in persons with aphasia (PWA) (Morris, Eccles, Ryan & Kneebone, 2017).

Anomia, the inability to retrieve words, is the most frequently reported, and often the most frustrating, symptom of aphasia. It commonly persists into the chronic stage (more than six months post-stroke) and constitutes a serious barrier to communication and functioning in daily life (Johansson, Carlsson & Sonnander, 2012). Although anomia, and aphasia in general, may respond to therapy many years after a stroke occurs, provision of specialist speech and language therapy is

far below that needed to provide optimal rehabilitation (Code & Heron, 2003). Currently, due to the underfunded state of the National Health Service, PWA typically receive just 12 hours of therapy, however, it is believed around 20 to 100 hours of therapy is necessary to improve language function (Brady et al, 2022; Clarke et al, 2018; Palmer, Witts & Chater, 2018; Bhogal, Teasell & Speechley, 2003). With an estimated 1.2 million stroke survivors in the UK, and an expected 27% rise in the number of people living with stroke by 2047, innovative intervention tools are required to expand the capacity for rehabilitation in order to provide the much needed increase in therapy dose (Stroke Statistics, 2017; Wafa et al, 2020).

One method for achieving this is by utilising digital technologies. The use of computer-based therapies in aphasia rehabilitation is a growing trend and a systematic review by Zheng, Lynch and Taylor (2016) concluded that these types of rehabilitation techniques can be as effective as traditional one-to-one therapy with a speech and language therapist. Furthermore, recently a large randomised control trial (Big CACTUS) demonstrated the efficacy of a clinician implemented, self-led computer therapy in improving word retrieval (Palmer et al, 2019). However, the evidence-base for truly automated digital intervention is sparse. Additional research is required to examine the effectiveness and acceptability of self-led therapy using different treatment techniques, and to elucidate the factors associated with positive outcomes.

Recent advancements in neuroimaging have provided a wealth of evidence showcasing the large and complex structural and functional architecture of the language network in healthy individuals (Hickok & Poeppel, 2004, 2007, 2015; Price, 2012; Fujii et al, 2016). However, the neural mechanisms underlying

recovery in the damaged brain are still poorly understood. Although studies investigating longitudinal changes following anomia therapy are the most commonly reported in the aphasia literature, inconsistent and contradictory results have been found. This is perhaps due to heterogeneity between PWA in terms of lesion characteristics and behavioural profiles, as well as methodological limitations such as small sample sizes, differing imaging approaches and variability in the treatment methods utilised (Kiran & Thompson, 2019; Doogan, Dignam, Copland & Leff, 2018). Further research, using both structural and functional imaging techniques, with larger populations and validated methods of analysis are required to understand the brain processes driving recovery. It is hoped that if the neural correlates underlying recovery can be identified, interventions can be designed in a way which boost neuroplasticity and improve treatment outcomes.

This thesis will contribute new evidence to the areas of anomia research described above by:

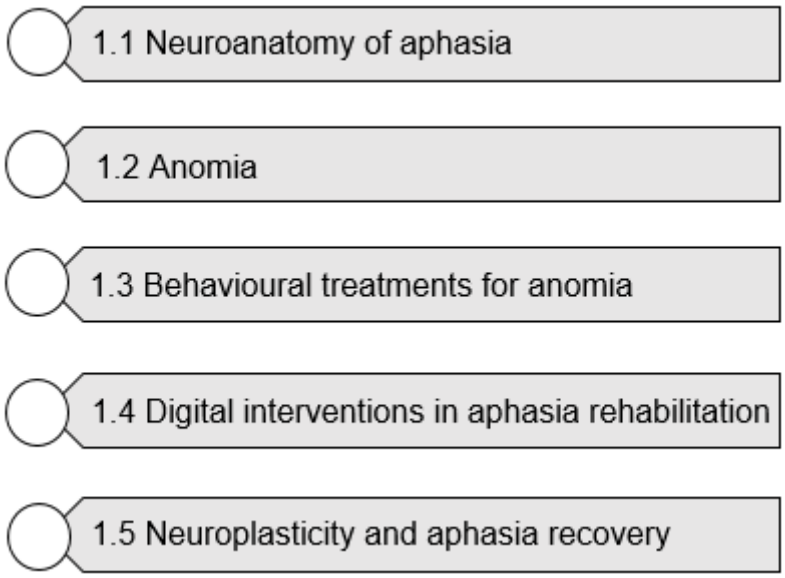
1.) Evaluating the efficacy of a self-led, tablet-based, anomia therapy ('iTalkBetter') in improving single word retrieval.

2.) Investigating whether demographic factors and performance on a battery of baseline neuropsychological tests can explain the variability in individuals' response to therapy.

3.) Identifying, using structural imaging, if there are treatment-induced changes in the structural integrity of the remaining cognitive and language networks.

4.) Examining, using functional imaging, if there are any therapy-driven effects in the language and cognitive networks.

In the following introduction, a background in the below areas will be provided:



1.1 Neuroanatomy of aphasia

Aphasia is a neurogenic language disorder which can affect any and all aspects of language and is characterised by morphological, semantic, phonological and syntactic processing impairments (Papathanisou, Coppens & Davidson, 2017). In people with stroke, it results from a focal brain lesion in the language dominant (usually left) hemisphere, typically involving the frontal, temporal and/or parietal lobes (Rasmussen & Milner, 1975; Szaflarski et al, 2002; Knecht et al, 2000; Price, 2010). The study of the neuroanatomy of language and aphasia started in the 19th century with Pierre Paul Broca's series of case studies with patients with localised brain lesions in the left frontal lobe (Broca, 1861a, 1861b, 1864). These lesions, in the pars triangularis (BA45) and pars opercularis (BA44) of the left inferior frontal gyrus (IFG) (Broca's area), were associated with severely non-fluent, agrammatic, speech production impairments. In 1874, Carl Wernicke expanded on Broca's work, identifying how lesions localised to the left posterior superior temporal gyrus (BA22) (Wernicke's area) resulted in disorders of comprehension with fluent but errorful speech and poor repetition (Wernicke, 1874).

Although these investigations led to the first theoretical model of language, which is still highly influential today, advancements in neuroimaging methods have highlighted the limitations of this classification framework. It is now accepted that the language system is comprised of a larger and more complex integration of multiple neural sites in the frontal, temporal and parietal lobes which function collectively to understand and produce speech (Price, 2012; Fujii et al, 2016). Currently, the most frequently reported neurological model of language is Hickok and Poeppel's dual stream account (Hickok, 2009; Hickok & Poeppel, 2004, 2007, 2015) (Figure 1.1).

This model describes two cortico-cortical pathways that form a ventral stream and a dorsal stream.

1.1.1 Dual stream model of language

The ventral stream supports sound-to-meaning mapping and therefore underlies auditory comprehension. The first stages of auditory analysis, specifically phoneme discrimination, occurs bilaterally in the superior temporal gyri (STG) (primary auditory cortex (Heschl's gyri) and planum temporale) and the superior temporal sulci (STS) (Saur et al, 2010; Okada et al, 2010; Vigneau et al, 2011; Turken & Dronkers, 2011; Penhune, Zatorre, MacDonald & Evans, 1996). The ventral stream then projects to a left-dominant system comprising of the middle temporal gyrus (MTG) and the inferior temporal gyrus (ITG) (where the phonological representation of the word is retrieved), and the anterior temporal lobe (ATL). The ATL is believed to be a hub which enables access to sematic representations, distributed throughout the cortex (Price, 2012). Damage to these grey matter areas and their underlying white matter tracts (the middle longitudinal fasciculus, the inferior longitudinal fasciculus and the inferior frontal-occipital fasciculus), results in disorders of speech comprehension (Hickok & Poeppel, 2015; Fernández Coello et al, 2013). This includes Wernicke's aphasia and transcortical sensory aphasia, a syndrome with poor auditory comprehension, fluent but errorful speech, and intact repetition.

Conversely, sound-to-articulation mapping, which translates phonological representations into articulatory processes to produce speech, is underpinned by the dorsal stream in a left-dominant network (Friederici, 2012; Hickok & Poeppel, 2007; Wise, 2003). The posterior part of the planum temporale, at the borders of the parietal and temporal lobes (Sylvian parietal-temporal junction, Spt), is the

sensorimotor interface of the dorsal stream which, although not speech specific, is densely connected to motor and auditory regions (Buchsbaum, Hickok & Humphries, 2001; Buchsbaum, Olsen, Koch & Berman, 2005; Hickok & Poeppel, 2015). From the Spt, the network projects to the IFG (Broca's area), the insula and the premotor cortex for speech articulation (Oh, Duerden & Pang, 2014). Lesions in these areas, and in the white matter tracts of the superior longitudinal and arcuate fasciculi, are associated with deficits of speech production, such as apraxia, Broca's aphasia and conduction aphasia (a disorder with a specific impairment in repetition) (Fridriksson et al, 2018; Price, Seghier & Leff, 2010; Price, 2010; De Smet, Paquier, Verhoeven & Marien, 2013; Fernández Coello et al, 2013). Whereas, widespread damage to both the dorsal and ventral streams results in global aphasia, a syndrome that severely affects comprehension and production of language.

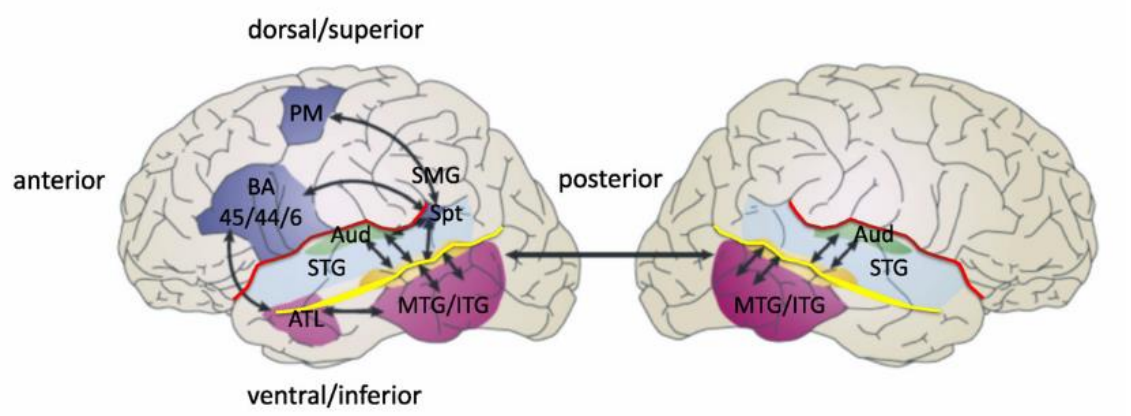


Figure 1.1 The dual-route model of language. Retrieved from Hickok (2009).

1.1.2 Neuroanatomy of speech production

The dual stream account provides an overview of language in the brain, however, functional imaging studies which have specifically investigated single word naming and propositional speech have provided further insights into the distributed network involved in speech production. Those employing single word naming paradigms

have attempted to separate word retrieval into phonological, semantic and syntactic levels of processing (Geranmayeh et al, 2012). These studies have suggested left temporal regions are involved in semantic selection; the retrieval of the corresponding phonological form then takes place in the left frontal cortex (Price, 1998), and later stages of production, such as generation of the articulatory code and programming of the articulatory sequence, are associated with the left IFG, the insula and the premotor cortex (Mechelli, Josephs, Lambon Ralph, McClelland & Price, 2007; Okada & Hickok, 2006; Indefrey & Levelt, 2004; Holland & Crinion, 2021).

Others have investigated the neuro-architecture of speech using discourse tasks to observe the entire speech production network and have identified an even wider range of implicated brain regions. For example, activations in the left inferior and medial frontal gyri, insula and superior parietal lobule have been consistently identified when producing connected speech (Haller, Radue, Erb, Grodd &Kircher, Price, 2012; Oh, Duerden & Pang, 2014). Geranmayeh et al (2014) also propose the involvement of bilateral domain-general cortical areas in supporting language output via top-down control processes, and include the fronto-parietal network (dorsolateral prefrontal cortex, precuneus, middle cingulate cortex, inferior parietal lobe and intraparietal sulcus) and the cingulo-opercular network (frontal operculum, dorsal anterior cingulate cortices, anterior insula, supplementary motor area and the anterior prefrontal cortex) (Hartwigsen & Saur, 2019; Price, 2012; Dosenbach et al, 2008). Furthermore, as the production of language relies on the same internal representations of knowledge and experience as comprehension does, overlap in activation for both perception and articulation in discourse have been identified in

areas which support episodic memory, such as the retrosplenial/ posterior cingulate cortex and the parahippocampus (Awad, Warren, Scott, Turkheimer & Wise, 2007).

1.2 Anomia

Word finding difficulties, or 'anomia', are a speech production impairment described as the hallmark of aphasia as they are the most prevalent, and often the most frustrating, symptom (Agostini et al, 2014; Goodglass & Wingfield, 1997). Anomia refers to a specific difficulty in retrieving words, even for common objects, and is characterised by the production of semantic, phonological and neologistic speech errors (paraphasias) and the use of non-specific words (such as 'thing') (Sze, Hameau, Warren & Best, 2021). Although anomia is an output impairment, recent research suggests it has no specific lesion location and is instead associated with an extensive cortical network which relies on the interaction between the ventral and the dorsal streams (Yourganov, Smith, Fridriksson & Rorden, 2015; Fridriksson et al, 2018; Hula et al, 2020). To say a word, one must first perceive a concept (for example, seeing an animal) and retrieve the semantic representation. Following this, the phonological features must be activated in order to programme the articulatory sequence and produce the word (Laine & Martin, 2006). Disruptions to any one (or a combination) of these components can result in difficulties with word retrieval, manifesting as paraphasias which can be analysed to identify the loci of breakdown.

1.2.1 Cognitive models of single word production

Cognitive models of single word production attempt to explain the retrieval of a word's semantic and phonological representations, as described above, using neuropsychological approaches. Although models tend to agree on the necessary

activation of meaning and sounds, differences exist in whether these activations are discrete or interactive.

Discrete models assume an independence of each step of word retrieval and are serial in nature. A classic cognitive neuropsychological discrete model of single word processing is displayed below in Figure 1.2 and is based on Patterson and Shewell's logogen model (1987) (Retrieved from Whitworth, Webster & Howard, 2014). For single word speech production, a speaker must first access the semantic system, then retrieve the phonological representation from the phonological output lexicon, assemble the phonemes and programme the articulatory sequence. Anomia, it is hypothesised, can result from damage to either the semantic system, the phonological output lexicon or phonological assembly, resulting in the common occurrence of semantic and/or phonological speech paraphasias. However, as these models are serial, they cannot account for other types of errors, such as mixed errors.

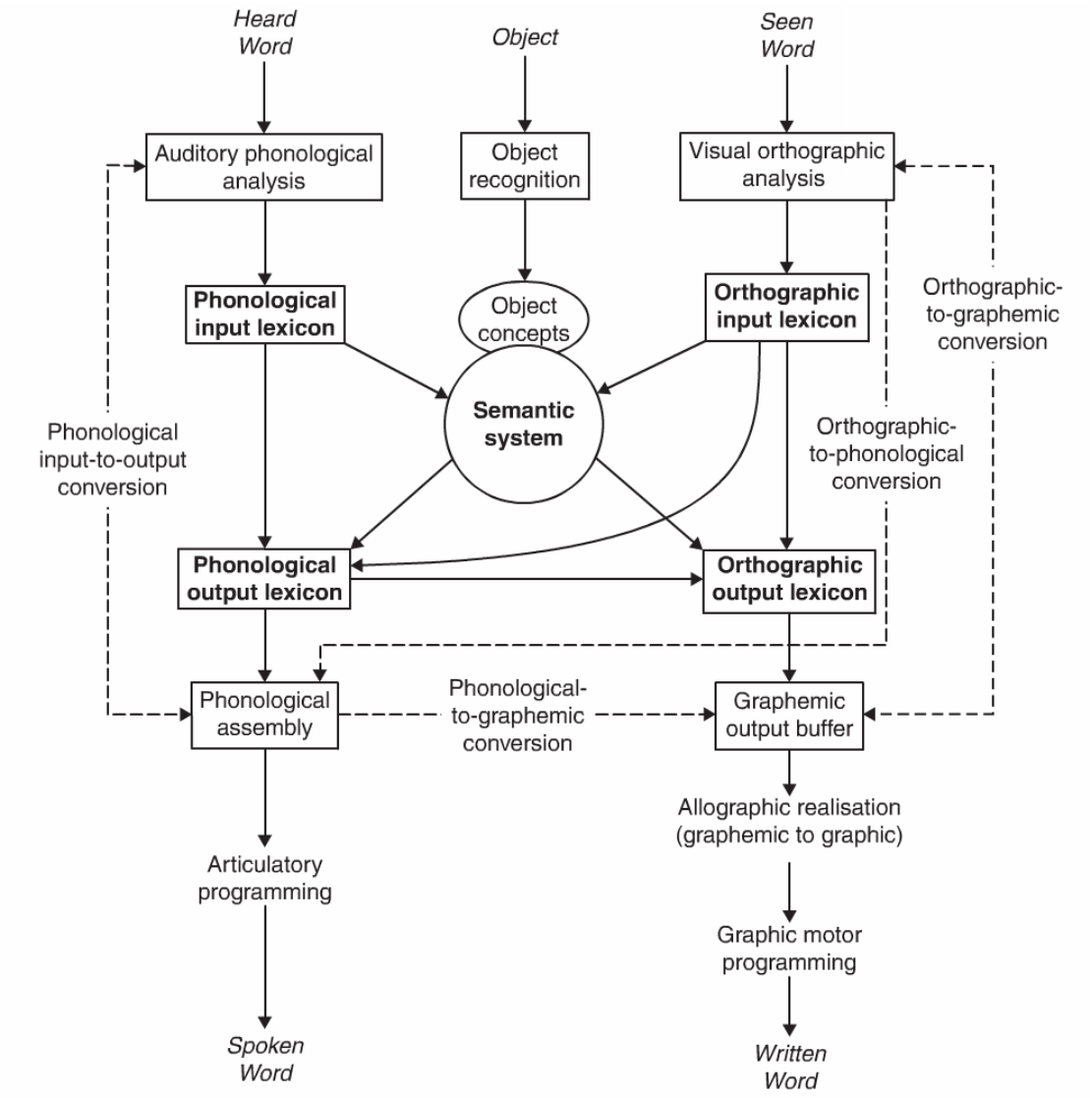


Figure 1.2 Cognitive neuropsychological model of single word processing. Retrieved from Whitworth, Webster and Howard (2014).

On the other hand, interactive models view word retrieval in a connected manner in which the stages interact with one another. As activation from one module spreads to another, activation also feeds backwards to keep the semantic representation, the lexical representation and the phonemes stable before the word is produced. One such model is the Semantic-Phonological interactive activation model by Dell, Schwartz, Martin, Saffran and Gagnon (1997) (Figure 1.3). In this model, the retrieval of words is dependent on semantic (S) and phonological (P) weights. (S) represents the strength of the connection between semantic and lexical layers,

whereas (P) is the connection strength between lexical and phonological layers. As in discrete models, paraphasias result due to breakdowns at different levels of retrieval. If (S) is weak, interference from competing semantically related words will result in semantic errors, however, if (P) is weak, phonological errors may occur (Oppenheim, Dell & Schwartz, 2010). Furthermore, as this model is bi-directional, it can also explain the presentation of other types of paraphasias, including unrelated and mixed errors, and neologisms which are thought to arise from subsequent phonological encoding impairments following incorrect word selection (Foygel & Dell, 2000).

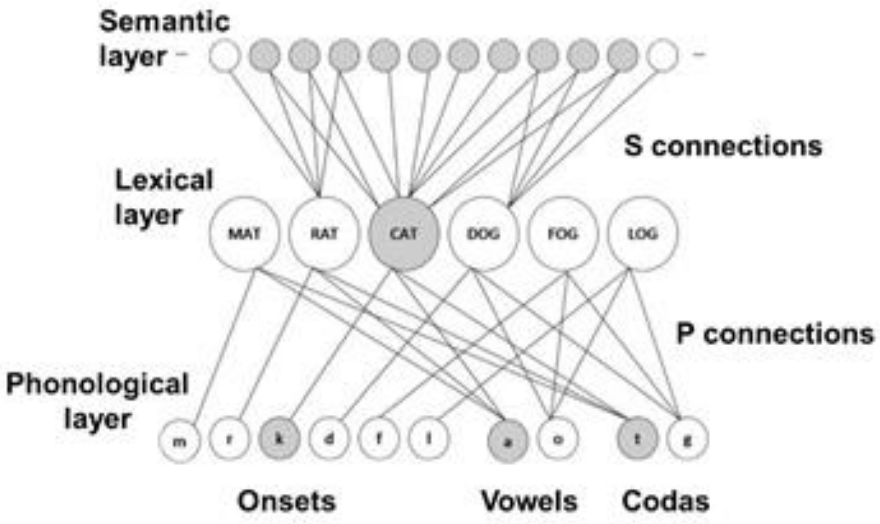


Figure 1.3 The Semantic-Phonological interactive activation model. Retrieved from Upton and Hope (2020).

Cognitive models do not intend to map language processes onto the structural architecture of the brain. However, differential neural signatures have been found for correct versus incorrect instances of word retrieval and the production of paraphasias in confrontation naming tests have been related to distinct areas of damage (Meinzer et al, 2013). For semantic errors, the regions identified relate to more ventrally located regions such as, the left ATL, the prefrontal cortex and the

MTG (Halai, Woollams & Lambon Ralph, 2018; Walker et al, 2011; Schwartz et al, 2009). Conversely, phonemic paraphasias and neologisms have been attributed to disruptions in the dorsal stream, including the premotor cortex, the left frontal lobe, the primary auditory cortex, and the arcuate and middle longitudinal fasciculi (Halai, Woollams & Lambon Ralph, 2018; Hula et al, 2020).

1.3 Behavioural treatments for anomia

Treatments for impairments of speech production are one of the most researched areas in aphasia therapy. Studies indicate intervention is efficacious in improving language outcomes, in comparison to no therapy, has short-term and long-term effects, and can benefit PWA even in the chronic stage (more than six months post-stroke) (Sze, Hameau, Warren & Best, 2021; Nickels, 2002; Breitenstein et al, 2017; Wisenburn & Mahoney, 2009; Raymer et al, 2008).

Behavioural interventions for word retrieval deficits generally fall into three main categories: social and environmental; compensatory; and impairment-based. These therapy approaches are not mutually exclusive and all may be required within a treatment programme, however, each approach employs differential strategies to support recovery. Social and environmental approaches tend to focus on aspects such as conversation in group therapy, training communication partners and adapting the environment to facilitate speech. Compensatory therapies aim to improve language outcomes by utilising relatively unaffected skills, such as gesture and writing (Greenwood, Grassly, Hickin & Best, 2010). Finally, impairment-based approaches strive to recover damaged language processes via re-learning techniques that aim to reinforce the production of target words by promoting neuroplasticity and reorganisation of disrupted language and cognitive networks

(Kleim, 2011). Although evidence does not necessarily, at the group level, indicate a superiority of one therapy technique over another (Brady, Kelly, Godwin, Enderby & Campbell, 2016), impairment-based interventions are the easiest to digitise which is why this treatment approach was selected in the current study. Therefore, for the purposes of my thesis, the following sections will discuss single-word, impairment-based therapies.

1.3.1 Single word anomia therapy

Impairment-based, single word therapies typically aim to improve word retrieval for a given selection of target words by employing semantic or phonologically-based approaches, alongside pictorial stimuli, which target the underlying lexical retrieval processes (Nickels, 2002; Davis & Pring, 1991). Semantic-based therapies attempt to strengthen conceptual representations and often focus on describing the properties of a target word. Alternatively, phonological-based interventions aim to strengthen phonemic representations and include repetition and phonemic or orthographic cueing. The descriptions and cues within these tasks can be generated by the individual to promote self-cueing (for example, in Semantic Feature Analysis (SFA) or Phonological Component Analysis (PCA)) or they can be provided by a therapist or computer programme, as in this study (Boyle, 2004, 2010; Leonard, Rochon & Laird, 2008; Van Hees, Angwin, McMahon & Copland, 2013; Sze, Hameau, Warren & Best, 2021).

1.3.1.1 Semantic and phonological cueing

Cueing aids in the retrieval of words by providing semantic or phonological information. Typical semantic cue types include providing the word category (for example, ‘pet’ for ‘dog’), or an associative word (for example ‘ringing’ for ‘telephone’),

and sentence completion. Phonological cues consist of providing either increasing or decreasing phonemic information (for example, ‘t’, ‘ten’ and ‘tennis’ for ‘tennis’), or words that rhyme with the target (Nickels, 2002; Nickels & Best, 1996; Python, Pellet Chevenal, Bonnans & Laganaro, 2021).

Although both phonological and semantic-based cues are efficacious, no single task has been found to be effective for all individuals. It has been predicted that differences in the success of these techniques depends on an individual’s loci of impairment, with those with phonological deficits benefitting more from phonological cues, and semantic cues supporting those with semantic impairments (Hicken, Best, Herbert, Howard & Osbourne, 2002; Neumann, 2018). Research indicates, however, that both semantic and phonological cueing can be effective for the same individual, with perhaps a superiority of phonological cueing benefitting those with all types of lexical deficits in anomia (Hillis 1998; Wamburgh et al, 2001; Li & Williams, 1991; Van Hees, Angwin, McMahon & Copland, 2013).

Indeed, in a within-subject study of 15 PWA by Python and colleagues (2021), the provision of semantic versus phonological cues in a picture naming task were compared to ascertain if the effectiveness of cue types corresponded to the level of impairment. Phonological cues were found to support word retrieval regardless of whether the participant had phonological, semantic or mixed phonological and semantic anomia. Contrastingly, the effectiveness of semantic cueing varied with anomia type; those with mixed anomia responded to both categorical and associative cues; only categorical cues assisted those with phonological anomia; and for semantic level deficits, no semantic cues were effective. However, this was a single session experimental paradigm, not a therapy study, so the lasting effects of cueing cannot be extrapolated. Furthermore, the participants had relatively mild

anomia, scoring between 71% and 97% on a shortened French version of the Boston Naming Test (Thuillard Colombo & Assal, 1992), making it difficult to generalise the results to a wider aphasic population.

Nevertheless, similar results have been identified in other experiments which have compared single sessions of semantic and phonological cueing, as well as in studies investigating the effectiveness of cue types in therapy programmes (Meteyard & Bose, 2018; Lorenz & Zeigler, 2009). In Lorenz and Zeigler's case series of 10 PWA, participants completed two blocks of therapy, consisting of around seven hours over two to three weeks, one of which employed phonological cueing in picture naming and the other utilising semantic cueing. As in the previous studies, phonological cues were found to be more effective than semantic cues for most participants. The authors do state, however, that improvements in picture naming following phonological cueing were not as long lasting as those observed after semantic cueing, although no long-term follow-up was conducted. There were also inconsistencies between the success of semantic and phonological cues and participants' underlying lexical impairment, so it is unclear which individuals responded best to which cue types.

It may be that differences between semantic and phonological cues are overstated and both can be equally effective, perhaps because many people with anomia have semantic and phonological deficits and most approaches tap into more than one level of lexical processing (Lorenz & Zeigler, 2009). There does seem to be a slight superiority for the majority of individuals, at least initially, for phonological cueing in picture naming therapy (Meteyard & Bose, 2018). This could be because in presenting a picture, semantic information about the object is given and then providing a phonemic cue facilitates the retrieval of phonological information.

Nonetheless, it is also worth noting that many studies investigating phonological cueing have provided orthographic cues within the therapy programme and, although not exclusively analysed, may be contributing to the success of phonological techniques (Sze, Hameau, Warren & Best, 2021; Hicken, Best, Herbert, Howard & Osbourne, 2002; Lorenz & Zeigler, 2009).

1.3.1.2 Errorful, errorless and error-reducing cueing

Errorful and errorless learning refers to two different learning techniques that can be applied within single word anomia therapy. Previously, the standard procedure in the provision of cues aforementioned was errorful learning, in which increasingly larger semantic, phonological and/ or orthographic cues are given in a hierarchical order following unsuccessful attempts at word retrieval (Nickels, 2002). In errorful programmes, the production of errors, be that paraphasias or no responses, is inevitable. Contrarily, errorless learning is an approach which aims to reduce the production of errors by the use of repetition. It is based on the view that errors should be minimised to diminish the chance of strengthening associations between an incorrect naming response and a target word (Baddeley & Wilson, 1994; Fillingham, Hodgson, Sage, & Lambon Ralph, 2003; Conroy & Scowcroft, 2012).

The effectiveness of errorful and errorless learning was investigated in a series of studies by Fillingham, Sage and Lambon Ralph (2005a, 2005b, 2006). In these research papers, the authors identified errorless learning to be as effective as traditional, errorful, hierarchical cueing and was better tolerated by participants. Other research has also identified comparable outcomes for errorless and errorful learning techniques, with both types significantly improving word retrieval, but differences between individuals in the acceptability of each approach (Conroy, Sage

& Lambon Ralph, 2009a). In their study, Conroy and colleagues gave nine anomic individuals two parallel therapy programmes. In the errorless condition, pictures were presented with the spoken and written target word a total of five times and participants had to repeat the word each time. In the errorful condition, increasing semantic and phonological cues were provided in a five-stage hierarchical design. The researchers found that, initially, all participants found the errorless therapy rewarding, but those with more mild deficits eventually found this condition needlessly repetitive and engaged more with the challenge of errorful therapy. On the other hand, the errorful condition was initially frustrating for all PWA but even more so for those with severe impairments who continued to prefer the errorless paradigm. Moreover, although personal preferences for each technique may differ, the role of feedback is important for both approaches and, if unable to retrieve a word, individuals should be provided with the correct response (Mckissock & Ward, 2007).

An alternative form of errorless learning which was put forward by Fillingham and colleagues (2003) is 'error reducing' learning. In this approach, decreasing semantic, phonological and/ or orthographic cues (as opposed to repetition alone as in 'pure' errorless learning) are provided following successful naming attempts to minimise errors but to also keep individuals engaged and challenged. In a study by Abel, Schultz, Radermacher, Willmes and Huber (2005) increasing cues (errorful) as opposed to decreasing cues (error reducing) were associated with greater improvements following therapy, though other research has found error reducing to be as effective as errorful (Conroy, Sage & Lambon Ralph, 2009b; Conroy & Scowcroft, 2012). Furthermore, PWA may prefer error reducing approaches as successful naming is more likely and those who increasingly improve are gradually

challenged, whereas those who are having difficulties continue to be supported, increasing motivation and longer term outcomes (Conroy, Sage & Lambon Ralph, 2009b). Due to this, and because of the positive outcomes associated with phonological cueing discussed in the previous section, an error-reducing phonological cueing paradigm was chosen to be implemented in the iTalkBetter therapy.

1.3.1.3 Generalisation

When evaluating the single word anomia intervention programmes discussed above, the generalisation of learning, in which improvements extend beyond the therapy context, is a key clinical outcome and should be the goal of all treatment programmes. This includes both within-level (from trained to untrained items) and across-level (from single words, to sentences, to conversation) (Webster, Whitworth & Morris, 2015). However, there is a paucity of research investigating generalisation and how, and indeed if, single word anomia therapy supports generalisation remains questionable.

Within-level generalisation (often considered 'response generalisation') refers to how improvements in trained items generalise to untrained items within the same linguistic level as the therapy (using the same stimuli), such as within a confrontation naming task (Webster, Whitworth & Morris, 2015; Thompson, 1989). There is evidence that, for bilingual or multi-lingual PWA, generalisation may occur across languages for words with similar linguistic roots (Ansaldo & Saidi, 2014). Additionally, training abstract words (low imageability words, for example, 'justice') may promote generalisation to untreated, contextually-related, concrete words (high imageability words, for example, 'judge') (Kiran, Sandberg & Abbot, 2009; Sandberg & Kiran,

2014) and generalisation between words could be more likely if they share similar features (Thompson & Shapiro, 2007). On the whole, however, improvements are generally item-specific in impairment-based therapy, with little or no generalisation to untrained items (Nickels, 2002, Wisenburg & Mahoney, 2009; Raymer et al, 2008).

Some research does propose that the efficacy of an intervention in promoting within-level generalisation depends on the approach utilised (Nickels & Best, 1996; Wisenburg & Mahoney, 2009; Webster, Whitworth & Morris, 2015; Middleton, Schwarts, Rawson, Traut & Verkuilen, 2016). In a meta-analysis of 32 studies by Sze, Hameau, Warren and Best (2021), semantic therapies were identified as the most successful approach for promoting improvements in untrained items, despite the fact they were not identified as a predictor for trained items. However, although 'semantic task' was the second most important variable for untrained items, the authors do discuss that this finding may be due to shared semantic similarities between trained and untrained items.

Other research suggests it is the type of lexical impairment that influences the generalisability of therapy. For example, Best and colleagues (2013) investigated the effects of an errorful phonological cueing paradigm in 16 PWA and found only those with more of a phonological deficit, as opposed to a semantic deficit, demonstrated some improvement in the production of untrained items. Furthermore, for those with phonological impairments, there is evidence that within-level generalisation may depend on both the therapy approach employed and the severity of the deficit. In a computer-based study comparing semantic and phonological therapy in two PWA, semantic therapy was more effective at promoting generalisation for the participant with a more severe phonological impairment and phonological therapy was superior for the participant who had a milder phonological deficit (Holland, John & Woollams,

2018). For those with semantic level impairments, it is unclear what techniques could be utilised to promote transference of skills to untrained items, and often PWA with a wider semantic deficit have poorer overall intervention outcomes, even for trained items (Dignam et al, 2016; Martin, Fink, Renvall & Laine, 2006).

Across-level generalisation (sometimes known as 'stimulus generalisation') refers to the ability to apply the skills learned within a specific treatment programme to other communicative contexts (using different stimuli), for example, improvements in the production of single words generalising to improvement in sentences and connected speech (Webster, Whitworth & Morris, 2015; Thompson, 1989). One may assume increases in the production of trained items following single word therapy would generalise to other linguistic contexts as the processes underlying retrieval are mechanistically the same (Herbert, Hickin, Howard, Osborne & Best, 2008). However, there is often a lack of generalisation to real-word communication following impairment-based anomia therapy, despite the implementation of tasks to promote this outcome (Nickels, 2002, Wisenburn & Mahoney 2009; Marshall et al, 2018).

It may be that the additional conceptualisation requirements for more complex speech contexts means improvements in single word retrieval are difficult to apply to connected speech. For example, in learning theories generalisation to other situations is dependent on the similarity to the task practiced and is more likely if the skills required are comparable to those in which the learning occurred (Osgood, 1949). Instances of generalisation to similar contexts are considered 'near transfer' and there may be additional steps which are necessary following single word anomia therapy to promote the application of learned skills across the communication continuum (Subedi, 2004). This theory has been suggested by Conroy, Sage and Lambon Ralph (2009c) in their study examining vocabulary production after single

word therapy across communication contexts with seven PWA. The greatest improvements seen were in a confrontation naming test (same linguistic level of therapy), followed by picture supported narratives (near transfer) and finally, unsupported narratives (far transfer). As there was a higher level of word retrieval within the picture supported narratives as opposed to the unsupported narratives, it implies that across-level generalisation can occur but is less likely for tasks that are dissimilar to those targeted in intervention. Furthermore, connected speech is not only more demanding in terms of conceptualisation requirements, it also more demanding in terms of speed; in fluent speech, around two words per second are produced. By employing methods that increasingly necessitate quicker word retrieval, in addition to improving accuracy, generalisation to connected speech may be promoted (Conroy, Sotiropoulou Drosopoulou, Humphreys, Halai & Lambon Ralph, 2018).

Although these studies provide some suggestions for promoting both within-level and across-level generalisation, the general consensus is that most impairment-based interventions do not promote generalisation. Learning is usually item-specific, with only around one in four individuals making significant gains in untrained items, which is still likely to be an overestimation due to repeated exposure to these items in testing paradigms (Nickels, 2002; Best et al, 2013). PWA also continue to face difficulties in real world communication, despite improvements made in therapy (Carragher, Conroy, Sage & Wilkinson, 2012). Due to this, clinicians and researchers should provide words in therapy that are functional and personally relevant to the individual and steps should be taken to promote the generalisation of these words across the communication continuum (Webster, Whitworth & Morris, 2015; Herbert, Best, Hickin, Howard & Osbourne, 2003; Conroy, Sage & Lambon Ralph, 2009c).

1.3.1.4 Word selection

In order to provide functional and personally relevant words, clinicians and researchers must work with PWA to individualise therapy and choose words that will maximise the impact of intervention (Palmer, Chater & Hughes, 2017). Typically, however, the words selected for practice in the majority of anomia therapy studies have been concrete, highly imageable, nouns and verbs. This is despite the fact the English language is made up of multiple word classes, including nouns, verbs, adjectives, adverbs, pronouns and prepositions, with differing levels of concreteness and imageability that are all critical to the ability to communicate effectively (Renvall, Nickels & Davidson, 2013; Binney, Zuckerman & Reilly, 2016).

Although there is a distinction between imageability and concreteness, with imageability referring to the degree of which a word conjures a mental image and concreteness/ abstractness referring to the extent of which a word can be experienced by the senses (Richardson, 1976), the terms are often used interchangeably in the aphasia literature. PWA frequently have increased impairments in the understanding and retrieval of abstract words, perhaps because these words have less specific, definite representations in the language network (Nickels & Howard, 1995; Kiran, Sandberg & Abbott, 2009; Newton & Barry, 1997; Crutch & Warrington, 2005). This leads to difficulties in expressing emotions and opinions even though research suggests these communication functions are a higher priority for PWA in comparison to communicating basic needs (Armstrong, 2005; Armstrong & Ulatowska, 2007; Worrall et al, 2011).

In standard practice, the creation of individualised words lists is based on personally chosen words via interviews with PWA and their communication partners,

observations, drawing on previously composed lists, and asking PWA to complete communication diaries (Beeke, Maxim & Wilkinson, 2007; Palmer, Chater & Hughes, 2017; Simmons-Mackie & Damico, 2001). However, there is a bias for PWA and their communication partners to choose concrete words and the question remains as to whether the words selected are truly relevant for people to convey what they want. To tackle this, Renvall, Nickels and Davidson (2013) propose two different approaches which should be combined for word selection in anomia treatment: personally chosen words and generally frequent words. Generally frequent words are those which are frequent across a population and, therefore, are likely to be useful for all speakers. One way in which to identify the most frequent words are by using language corpora, which Renvall and colleagues did in their 2013 study. In their research paper, the authors compared two language corpora which were composed using spoken sources, the CELEX (1.3 million words) and the SUBTLEX-US (51 million words) databases (Baayen, Piepenbrock & Gulikers, 1995; Brysbart & New, 2009). They found that in the 1000 most frequent words, concrete nouns and verbs were prominent but so were abstract nouns and verbs. Furthermore, adjectives (describing living and non-living things, and emotions) were the third biggest word group, followed by adverbs (words which modify a verb, adjective, another adverb or a word group). These findings indicate generally frequent words are less imageable and concrete than items typically targeted in therapy.

As a follow-up to this research, Renvall and Nickels (2019) completed a single case treatment study using a traditional therapy technique (repetition in the presence of a picture (errorless learning)), to improve the production of abstract emotive adjectives. This was to assess whether standard word retrieval therapy could be used in the treatment of the more abstract items identified in the generally frequent word lists. In

the therapy, the participant, GEC, trained on 72 items which consisted of 36 positive emotive adjectives (for example, ‘happy’, ‘hilarious’) and 36 negative emotive adjectives (for example, ‘stupid’, ‘boring’), over a two week period, completing eight hours of therapy. Following therapy, GEC significantly improved in the production of positive emotive items and this improvement was maintained 11 weeks after the cessation of therapy. The authors also note there was an improvement in the production of negative items, however, this was not significant.

The above research highlights the importance of selecting a myriad of word types, of both personally chosen and generally frequent words, within anomia therapy to provide a treatment programme that will bring the most benefit to PWA. For generally frequent words, published and accessible language corpora may be useful sources of information to select words for therapy. Although the vast SUBTLEX-US database was based on American film and television series subtitles, van Heuven, Mandera, Keuleers and Brysbaert (2014) have now created a SUBTLEX-UK database for British English which has been utilised in the current study. Moreover, Renvall and Nickels’ case study provides preliminary evidence that standard therapy may be effective for improving abstract words, although more research is required. As PWA do exhibit a ‘concreteness’ effect, in which concrete, highly imageable, words are easier to retrieve than abstract words and require less cueing, studies also need to examine how much therapy is required to improve these more difficult items (Walker & Hulme, 1999; Conroy, Snell, Sage & Lambon Ralph, 2012; Nickels & Howard, 1995; Kiran, Sandberg & Abbott, 2009).

1.3.1.5 Dose and intensity

A crucial and important aspect of single word anomia rehabilitation, regardless of the type of treatment selected (phonological or semantic, errorful or errorless), is repeated practice. Practice enhances skill acquisition, consolidates learning and can support improvements in word retrieval long-term (Harnish et al, 2018). Although repeated practice is necessary in impairment-based therapy, two important measures should be considered when implementing treatment programmes; dose (how much therapy, usually in hours) and intensity (frequency of therapy). A number of reviews and meta-analyses have provided evidence for the benefits of high dose aphasia therapy, suggesting that for a significant and meaningful improvement in language, around 20-100 hours of therapy is required (Bhogal et al, 2003; Brady et al, 2016; Brady et al, 2022; Breitenstein et al., 2017; Cherney, Patterson, & Raymer, 2011; Doogan, Dignam, Copland & Leff, 2018; Harvey, Carragher, Walsh Dickey, Pierce & Rose, 2022).

Currently, however, it is unclear whether high intensity schedules of massed practice or low intensity schedules of spaced practice are more beneficial in improving outcomes and what exactly constitutes to a 'high' or 'low' schedule (Sage, Snell & Lambon Ralph, 2011; Dignam, Rodriguez & Copland, 2016). Some define high intensity schedules as those which consist of five hours or more a week, while others consider this a low intensity schedule (Stahl et al, 2018; Dignam et al, 2015). Furthermore, it is often difficult to disentangle dose from intensity in the aphasia literature as many studies, including that by Bhogal and Brady's Cochrane review, confound the two, with studies employing high intensity schedules also reporting higher doses than low intensity schedules.

Research that has attempted to control the two measures have produced inconclusive results. For example, Dignam and colleagues (2015) found no difference between a high intensity programme (16 hours per week, over three weeks) and a low intensity programme (six hours per week, over eight weeks), whereas as Breitenstein et al's (2017) large randomised control study with 158 participants did find an overall benefit of a high intensity, massed practice schedule (≥ 10 hours per week, over three weeks). Although this provides some evidence favouring massed practice, the research is limited and almost all studies have found a positive significant impact of therapies which are given at a higher dose, regardless of what intensity schedule is employed (Bhogal, Teasell & Speechley, 2003; Basso, 2005; Brady et al, 2016). Additionally, massed practice schedules are often associated with significant drop out rates, illustrating that intensive schedules are not suitable for all PWA (Brady et al, 2016).

In order to tease apart the effects of dose and intensity, and to standardise the terminology across intervention studies to enable more valid comparisons, Harvey, Carragher, Walsh Dickey, Pierce and Rose (2022), propose a dose-intensity framework that comprises of the following definitions: 'therapeutic element': the therapeutic input; 'session dose': how many minutes or therapeutic elements are provided in a session; 'session intensity': the rate of which therapeutic elements are given; 'session frequency': the number of sessions per week; and 'total dose': the overall amount of therapy given. In a single word anomia therapy programme, a therapeutic element may be a list of target words and session dose could refer to the individual dose or repeated practice of each target word ('exposure'). There is research which suggests providing more words leads to enhanced outcomes (more therapeutic elements), however, a greater session dose (greater number of

repetitions of, and exposure to, target words) may lead to better overall improvement (Thomas, Lander, Cox & Romani, 2020; Snell, Sage, & Lambon Ralph, 2010; Off, Griffin, Spencer & Rogers, 2016; Godecke et al, 2013). Clearly defining dose and intensity terms at both the session and overall level will provide more guidance to researchers and clinicians when describing and reporting on aphasia therapy.

Overall, dose and intensity research is a relatively new field in aphasia therapy and considerably more work is necessary before conclusions can be drawn to inform best practice. There needs to be clear definitions in the literature so the outcomes of intervention studies can be evaluated and compared. Systematic investigations of therapy intensity, without manipulations in dose, is warranted and researchers should examine both person-related factors (for example, aphasia severity, time since stroke, age) and intervention-related factors (for example, intervention type and schedule) to assess what type, how much and how frequent therapy needs to be for each PWA (Brady et al, 2022). This could be achieved through the use of adaptive trial designs that use dose-response modelling to identify optimal dose and intensity schedules on an individual basis, rather than employing one approach for all (Doogan, Dignam, Copland & Leff, 2018; Bhatt & Mehta, 2016).

1.3.2 Explaining and predicting response to therapy

Although many studies have demonstrated the effectiveness of single word anomia therapy, and researchers have attempted to elucidate the factors underlying treatment response, no variables have been identified which consistently predict outcomes. Within the same intervention programme, there can be considerable variability in individuals' therapeutic gains and even those with similar linguistic and non-linguistic profiles show differential improvement patterns (Wisenburn &

Mahoney, 2009; Nickels, 2002). It has been hypothesised that demographic factors, language and cognitive factors, lesion characteristics, and functional network activation may offer value in predicting treatment response.

The main demographic factors discussed in the literature which are thought to impact therapy outcomes are age and time post-stroke. Age has been suggested to play an important part in treatment-related outcomes, with younger PWA showing superior response to therapy (El Hachioui et al, 2013; van de Sandt-Koenderman et al, 2008), however, opposing evidence disputes the usefulness of age as an explanatory variable (Pompon et al, 2017; Code, Torney, Gildea-Howardine & Williams, 2010). Additionally, some propose providing treatment at an earlier stage in post-stroke recovery leads to better outcomes (Kirmess & Maher, 2010), but those who are in the chronic stage of aphasia can still make significant gains in therapy (Palmer et al, 2019; Brady et al, 2016; Raymer et al, 2008; Meinzer et al, 2004).

There is a similar conflicting research-base for language factors, specifically baseline severity, in predicting therapy response. Some studies have identified significant correlations between baseline severity and language outcomes, with those who are milder improving more, both immediately and at longer term follow-ups (Lambon Ralph, Snell, Fillingham Conroy & Sage, 2010; Kiran, 2016). Others have found those who have more severe deficits do better (Laska et al, 2001; Robey, 1998), and still more have discovered no link between baseline severity and treatment response (Code et al, 2010; Persad, Wozniak & Kostopoulous, 2013; Pompon et al, 2017; Hope et al, 2021).

Although not as well researched, there does seem to be more consistent findings for the role of cognition in therapy response, though exactly which areas of cognition are

important is still under investigation. Better executive functioning has been linked with improved naming accuracy (at post-therapy and at follow-up), perhaps due to a greater ability to learn and apply naming strategies (Yeung & Law, 2010). Furthermore, Lambon Ralph and colleagues (2010), pooled data from four comparable studies to identify both language and cognitive factors that explained response to naming therapies for 33 PWA. Again, executive function significantly positively correlated with treatment outcomes, along with attention and visuo-spatial memory. On the other hand, Dignam and colleagues (2017) found evidence suggesting verbal short-term memory was the strongest predictor for response to anomia therapy. In addition to general cognition, psychosocial factors such as social support (social network size and relationship status) and mood (anxiety and depression) can also influence engagement with therapy (Worrall, Hudson, Khan, Ryan & Simmons-Mackie, 2016).

A growing body of research also indicates the prognostic value in the use of brain data to explain and predict response to aphasia therapy (Watilda & Balarabe, 2015; Plowman, Hentz & Ellis, 2012; Kiran & Thompson, 2019). This includes the lesion characteristics of size and location, and, although lesion location seems to be a more reliable predictor (Hope et al, 2021; Price, Seghier & Leff, 2010; Plowman et al, 2012), the importance of pre-therapy brain structures in anomia recovery differs from study to study. For example, Marcotte and colleagues (2012) found increased damage to Broca's area (BA45) was associated with poorer response to single word intervention, whereas in a study by Fridriksson, Bonilha, Baker, Moser and Rorden (2010), damage to a region in the left temporo-occipital area was linked with a lesser ability to improve naming following therapy. The integrity of pre-therapy functional networks, in both the right and the left hemispheres, may also help to predict

treatment outcomes. There is evidence that the bilateral integrity of the memory network and the integrity of the left language network can predict response to therapy both immediately and at longer term follow-up (Menke et al, 2009), and that pre-treatment activity in different regions in the left hemisphere correlate with treatment success depending on the task utilised (semantic or phonological) (Van Hees, McMahon, Angwin, de Zubicaray, & Copland, 2014).

The variability in predicating therapy response across PWA (as highlighted in the above studies) is confounded by differential treatment programmes and the use of varied assessment methodologies, complicating the identification of explaining and predicting factors. The use of neuroimaging may offer an additional prognosis power, over and above that provided by behavioural and demographic data alone (Hope et al, 2021; Iorga et al, 2021). However, due to large heterogeneity in brain structure and function across PWA, researchers should consider methods which combine multiple variables. For example, Aguilar and colleagues (2018) found that a model which made use of both behavioural and brain data had better explanatory and predictive power for single word reading therapy response than those that relied on just one variable group. Future research is required to build upon this evidence using larger data sets and investigating response prediction across different intervention programmes to illuminate who is likely to benefit from a particular therapy programme, and how treatments can be tailored for the individual.

1.4 Digital interventions in aphasia rehabilitation

Digital interventions (apps and web-based therapies) have grown in popularity over recent years and there is evidence that the utilisation of technology in aphasia rehabilitation is an effective tool (Brady et al, 2016; Zheng, Taylor & Lynch, 2016;

Des Roches & Kiran, 2017). Many applications deliver speech and language therapy over internet video conferencing and these can provide similar improvement outcomes as face-to-face therapy (Simic et al, 2016; Harnish et al, 2014; Woolf et al, 2016). However, these types of therapy approaches (as well as traditional face-to-face intervention) require many hours of therapist input and are expensive, meaning they are often not accessible to many PWA. The real benefit of harnessing technology in aphasia intervention is to create automated, self-led therapies that PWA can use at home without the need for clinical input. This provides a cost-effective and convenient approach to delivering the much needed high dose treatment programmes necessary to improve language recovery.

Computerised self-led therapies (CSLT) offer a way for individuals to access high dose therapy in addition to their usual care and give PWA some autonomy over their language recovery (Godlove, Anantha, Advani, Des Roches & Kiran, 2019). Additionally, they are a particularly useful resource for those who live in remote areas where access to speech and language therapy is difficult, and real-world research has shown usage of digital apps is higher in areas with limited rehabilitation services (Munsell, De Oliviera, Saxena, Godlove & Kiran, 2020). Studies have also identified that PWA who have no, or very limited, experience with computers are able to navigate and use digital therapies independently and the adoption of such technology is not affected by typical barriers such as age (Ramsberger & Marie, 2007; Munsell et al, 2020).

For single word anomia intervention specifically, a handful of smaller studies have demonstrated the effectiveness of CSLT in improving word retrieval (Routheir, Bier & Macoir, 2016; Mason et al, 2011). However, the largest study, and the first multi-centred, randomised control trial, which has investigated CSLT for word finding

difficulties is the Big CACTUS research project (Palmer et al, 2019). 278 participants were recruited into this study, of which 83 completed the experimental therapy arm, using the StepByStep© programme (Palmer et al, 2012), targeting the production of 100 personally relevant words over a six-month period. Following therapy, participants in the experimental arm made large and significant improvements in the production of trained items and, although this improvement did not generalise to conversation, gains were maintained six months after treatment. This important study provides evidence on a large scale that CSLT is feasible and effective.

One point which must be noted in the Big CACTUS study is that a therapist was still required to set up the treatment programme and throughout the six month intervention period, participants were visited once a month by an assistant or volunteer. Although this is still limited contact time in comparison to traditional and video conferencing approaches, truly automated therapies would provide clinicians and PWA freedom from this process. This would then enable face-to-face sessions to focus on language tasks which are more difficult to complete independently, such as those targeting functional communication. It would also be interesting to know if participants would have completed the necessary dose to improve their word finding abilities without regular visits from the research team. Even in tightly controlled research studies which have employed CSLT in other language domains, large variations in dose have been observed across individuals (Woodhead et al, 2017; Fleming et al, 2020).

Overall, technology in aphasia rehabilitation is a growing trend and there is evidence it is effective, achievable and cost-effective. CSLT in particular offers access to high dose therapy that PWA often do not receive as part of their usual care. It also provides researchers and clinicians with the ability to monitor aspects such as dose

and intensity, and keeps intervention programmes consistent across individuals, enabling proper comparisons between therapy techniques (Dignam et al, 2016). Additionally, the large amounts of data that can be collected through self-led digital therapies could be used in the future to individualise treatment programmes and make evidence-based decisions (Cordella et al, 2022).

1.5 Neuroplasticity and aphasia recovery

The advancements in neuroimaging in the last few decades have demonstrated the brain's capacity to adapt and reorganise in response to the environment, encoding novel experiences and learning new behaviours (Kleim & Jones, 2008). This experience-dependent plasticity has been shown in structural and functional imaging studies across multiple domains in healthy individuals. In structural imaging, greater regional volume in grey matter (GM) and white matter (WM) has been correlated with differing levels of expertise and superior proficiency in tasks such as second language learning and phonetic transcription (Mechelli et al, 2004; Golestani, Price & Scott, 2011). Longitudinal structural imaging has also revealed increases in volume in response to learning new skills, such as foreign language acquisition, juggling and playing golf (Mårtensson et al, 2012; Draganski et al, 2004; Bezzola, Merillat, Gaser & Jancke, 2011). In functional imaging, differences in neural activity in the expert brain have been found for ballet dancers and racquet ball players (Bangert & Schlaug, 2006; Pearce, Thickbroom, Byrnes & Mastaglia, 2000), and longitudinal changes have been observed following musical practice, auditory training, artificial language acquisition and foreign language learning (Herdener et al, 2010; Tremblay & Kraus, 2002; Xue, Chen, Jin & Ding, 2006; Barbeau et al, 2017).

Importantly, in aphasia research, neuroplasticity is also the mechanism by which the damaged brain spontaneously recovers function following a stroke and re-learns lost behaviours in response to therapy. A number of studies have investigated neuroplasticity in aphasia, whether spontaneous or related to treatment, but the brain processes underlying recovery remain elusive. It is hoped that by understanding these neural correlates, interventions can be designed in a way which boost neuroplasticity and improve treatment outcomes. The following sections will provide an overview of the current research for neuroplasticity in aphasia and will detail key evidence for treatment-induced structural and functional change following anomia therapy.

1.5.1 Spontaneous change

It was previously believed that language improvements following stroke plateaued in the chronic stage of recovery (more than six months post-stroke), with no further reorganisation or change in the structural architecture or functional networks of the brain (Culton, 1969; Sarno & Levita, 1971). But, as discussed in the above sections, even many years after a stroke occurs, behavioural change, and therefore neuroplasticity, is possible (Kiran & Thompson, 2019). The precise course of brain recovery, however, is complicated by the neurobiological sequela of stroke which persists into the chronic stage, leading to further changes in structure and function over time. This includes both local dysfunction at the lesion site and also damage to areas which are functionally and/or anatomically connected to the region of damage (Hartwigsen & Saur, 2019). These secondary disruptions are caused by diaschisis, distant neurophysiological changes due to deafferentation of neurons connected to the lesion site (Brodtmann et al, 2020), and Wallerian degeneration, degeneration of an axon of a damaged nerve fibre remote from the lesion (Zhang, Zhang, Xing, Liang

& Zeng, 2012). Overall, these changes lead to lesion expansion and gradual brain shrinkage years after a stroke occurs, over and above that expected in normal aging (Seghier, Ramsden, Lim, Leff & Price, 2014; Brodtmann, 2021).

Neuroplasticity is additionally confounded by heterogeneity across participants in factors such as lesion characteristics, behavioural and cognitive profiles, and demographics such as age (Li, Mukadam & Kiran, 2022; Kiran & Thompson, 2019). This, and variability in imaging techniques, has led to conflicting findings in how brain changes correlate with recovery. For example, some research suggests spared areas in the left hemisphere (LH) are vital for recovery (Saur et al, 2006; Perani et al, 2003; Meier, Kapse & Kiran, 2016; Szaflarski, Allendorfer, Banks, Vannest & Holland, 2013; Allendorfer, Kissela, Holland & Szaflarski, 2012; Szaflarski et al, 2011; Turkeltaub, Messing, Norise & Hamilton, 2011). Others report the importance of recruiting language homologue regions in the right hemisphere (RH) (Xing et al, 2016; Skipper-Kallal, Lace, Xing & Turkeltaub, 2017; Rosen et al, 2010; Fridriksson, Baker & Moser, 2009; Crinion & Price, 2005; Leff et al, 2002) and several suggest bilateral recruitment is necessary (Cao et al, 1999; van Oers et al, 2010; Griffis et al, 2017). The role of the RH in recovery may not be surprising considering that across many language tasks healthy individuals display bilateral activations (Vigneau et al, 2011), and RH homologues are structurally and functionally connected to the LH language network (Turken & Dronkers, 2011).

The recruitment of RH homologues, either independently or in addition to spared LH regions, however, remains a hotly debated subject in the aphasia literature, with some suggesting the RH is supportive of language recovery, and others suggesting it is inhibitory (Price & Crinion, 2005; Winhuisen et al, 2005; Cocquyt et al, 2017). It is likely the RH plays a differential role across individuals as significant variability has

been found in the lateralisation of language networks in healthy populations, suggesting hemispheric contributions to language processing is heterogeneous (Catani et al, 2007). Furthermore, in language recovery following stroke, spontaneous structural changes (hypertrophy) in the right hemisphere have been found to correlate with and predict both improvements and declines in language performance in the chronic stage (Hope et al, 2017).

Brain regions associated with the domain-general networks also seem to play an important part in aphasia recovery. Although the domain-specific network outlined in section 1.1 underlies language processing, this network interacts with bilateral domain-general networks involved in processes such as attention, working memory and cognitive control (Hartwigsen & Saur, 2019). These domain-general networks can be subdivided into a fronto-parietal network (central executive network), including the dorsolateral prefrontal cortex (DLPFC), precuneus, middle cingulate cortex, inferior parietal lobe and intraparietal sulcus; and a cingulo-opercular network (salience network) comprising of the frontal operculum, dorsal anterior cingulate cortices (dACC), anterior insula, supplementary motor area and the anterior prefrontal cortex (aPFC) (Dosenbach et al, 2008; Stockert et al, 2020; Seeley, 2019). The exact role of these networks within the language system remains under investigation and the precise function of some of the regions is still unclear. For example, the aPFC (BA10) is one of the largest brain areas in humans but its function is complex and poorly understood. It has been hypothesised to act as a gateway to cognitive processing, involved in working and episodic memory, multiple task co-ordination, attention, awareness of competence, and inhibition (Burgess & Wu, 2013; Burgess, Dumontheil & Gilbert, 2007; Gilbert et al, 2006; Fleming, Weil, Nagy, Dolan & Rees, 2010; de Zubicaray, Zelaya, Andrew, Williams & Bullmore,

2000), and may be connected, via the extreme capsule, to the STS and the insula (Petrides & Pandya, 2007).

Although further research is required to understand the role of the domain-general networks, in terms of both cognition and their role in language, recruitment of these broader, functional brain systems have been found to facilitate recovery, particularly for those with extensive damage in the language network (Brownsett, Warren, Geranmayeh, Woodhead, Leach & Wise, 2014; Geranmayeh, Brownsett & Wise, 2014; Kiran, Meier & Johnson, 2019). For example, in a study by Stockert and colleagues (2020), 34 participants with either left frontal (n=17) or left temporo-parietal (n=17) damage were recruited into a longitudinal study to assess patterns of language-related brain activation from the acute to the chronic stage of recovery. The researchers found that, for those with larger lesions, stronger activation in the chronic stage in right DLPFC was associated with greater improvements. Additionally, for those with left temporo-parietal damage, increased activation in the left insula correlated with behavioural improvements, indicating an additional factor of the effect of lesion site on the role of domain-general networks in recovery.

1.5.2 Treatment-induced neuroplasticity

1.5.2.1 Structural reorganisation

Only a small number of studies have explored treatment-induced longitudinal structural changes in aphasia recovery, likely due to the difficulty in capturing subtle microscopic changes and differences in processing pipelines, leading to a lack of replicability between studies (Thomas & Baker, 2013). Additionally, sample size affects the reproducibility of results, with small sample sizes both over and under estimating effect sizes, and large sample sizes yielding highly significant results even

when the effect sizes are negligible (Lorca-Puls et al, 2018). Two of the studies that have investigated structural changes used diffusion tensor imaging (DTI) to investigate adaptation in WM. DTI is a magnetic resonance imaging (MRI) technique based on the anisotropic properties of tissue, measuring the microstructure of the brain using the variation in diffusion of water molecules (O'Donnell & Westin, 2011). A commonly used measure within DTI is fractional anisotropy (FA), where greater diffusion, and therefore better defined pathways (white matter integrity), is represented by lower FA values.

In the first study, Schlaug, Marchina and Norton (2009) investigated white matter integrity before and after Melodic Intonation Therapy (MIT) in six participants with chronic Broca's aphasia. Using DTI, they found significant increases in the number of fibres and in the volume of the right arcuate fasciculus, which, in the LH, is a key white matter tract underlying the ventral stream in language processing. However, although the authors report the participants made significant improvements in speech production (picture naming, picture description and conversation) following therapy, they did not find a direct relation between behavioural change and WM reorganisation, perhaps due to the small number of participants included in the study. In the second study, Wan and colleagues (2014) also used MIT in a group of 11 participants with non-fluent aphasia. At post-therapy, reductions in FA were identified in the RH in the IFG and in the posterior part of the STG (pSTG), and the changes in the right IFG were associated with improvements in speech fluency, supporting the role of the RH in recovery.

Another study has built on these findings of changes in WM following speech production therapy by examining reorganisation of residual GM using a novel diffusion MRI approach to quantify non-Gaussian water diffusion properties, known

as diffusional kurtosis imaging (DKI). In this retrospective study of 33 participants, Chang and colleagues (2021) observed that increased microstructural integrity in the left MTG and the left ITG correlated with improved picture naming and was associated with decreases in semantic errors post-therapy. However, the treatment provided to these participants was a receptive language task and there was not a statistically significant improvement in naming. Additionally, the study design included a transcranial direct current stimulation (tDCS) arm and, although tDCS was included as a predictor of no importance in the model, it is unclear whether this influenced changes.

No study, to my knowledge, has employed longitudinal structural MRI imaging to examine changes in both GM and WM following speech production treatment. However, a study by Fleming and colleagues (2020) used this imaging method to investigate structural changes from pre- to post-therapy following a CSLT of spoken word comprehension. Following 12 weeks of therapy (average of 85 hours), participants significantly improved their comprehension of trained words and this gain was associated with change in the volume of GM in the pSTG of the RH and change in the WM of the middle STG in the LH. Although not a speech production study, this finding also adds to the evidence suggesting a supportive role of the RH in language recovery. Furthermore, the within-subject design of this study, whereby participants acted as their own control, comparing a no therapy period to a therapy period, avoided confounding variability due to between-subject variance and took account of changes related to time (for example, due to aging and general atrophy).

1.5.2.2 Functional reorganisation

The majority of studies investigating treatment-induced neuroplasticity in aphasic populations have employed functional MRI (fMRI) and, as in spontaneous recovery, have identified inconsistent results (Kiran & Thompson, 2019; Crinion & Leff, 2015). In a review by Thompson (2017) of fMRI findings between 1996 and 2016 from 41 studies examining therapy-induced change, it was found that studies have reported changes in activation in intact tissue in the LH, in the RH and bilaterally. Of the 628 PWA included in the studies, 99 participants showed upregulation of neural activation in the left hemisphere, 90 showed recruitment of neural tissue in the RH, and 439 participants showed increased activity bilaterally. Furthermore, a number of studies have also identified decreases in activation which were associated with positive therapeutic outcomes.

The association between intact regions in the LH and improvement following anomia therapy has been shown in several studies. For example, Rochon and colleagues (2010) completed PCA therapy with two aphasic individuals and compared their pre- and post-therapy activation with that of two PWA who did not complete therapy and 10 healthy controls. The two participants who received treatment improved their naming and in the post-therapy fMRI task (semantic and phonological judgement), significant areas of increased activation were identified which correlated with improvements in the left frontal and temporal regions. One out of the two PWA control participants also showed a change in activation on the second scan but this was not associated with gains in naming, and the other healthy control participants showed stable activation patterns over the two scans. This implies the increased LH activity for the therapy participants contributed to improvements post-treatment. However, this study included only a small sample size and there were large

variations in scan intervals between the participants. Also, the fMRI tasks did not involve overt naming so changes in neural activity may not have reflected changes in speech production.

In a larger study by Fridriksson and colleagues (2010), 19 participants completed 30 hours of anomia therapy over a two-week period. The therapy consisted of errorful cueing hierarchies, with half of the words targeted using semantic-based cueing, and the other half undergoing phonologically-based cueing. Multiple fMRI scans were completed before, during and after therapy, and the task consisted of naming trained items from the therapy content. Both treatment approaches resulted in statistically significant improvements in naming and these gains were associated with increased brain activation in anterior (premotor, including superior portion of Broca's area) and posterior (parietal lobe) areas of the LH. Similar findings were also identified by Marcotte and Ansaldo in 2010 and Marcotte and colleagues in 2012 following SFA therapy. In the study 2012 (which included nine PWA), increased accuracy in naming was associated with greater activation in the left precentral gyrus (primary motor cortex) at pre-therapy, and recruitment of the left inferior parietal lobule post-therapy. Additionally, successful responders recruited fewer brain regions post-therapy, with no further activations in the RH, indicating that the LH was associated with superior recovery.

Although the above studies provide evidence for the importance of the LH in recovery, others have found that the recruitment of the RH, including both left language homologues and domain-general areas, is also essential for therapy driven functional change. For example, Raboyeau and colleagues (2008) completed a word finding therapy with 10 PWA and found increased activity in the RH, specifically in the IFG and the insula, correlated with improved lexical retrieval post-intervention.

On the other hand, in a study by Johnson, Meier, Pan and Kiran (2019), 26 participants with chronic aphasia completed 12 weeks of semantic naming treatment and the recruitment of intact LH language regions and RH homologues was associated with recovery. Pre- and post-therapy activation patterns were investigated using an fMRI picture naming paradigm and those who responded favourably to treatment had increased neural activity bilaterally in the inferior frontal gyri and in the right middle frontal gyrus, relative to 17 healthy controls.

Additionally, in a study by van Hees, McMahon, Angwin, de Zubicaray and Copland (2014) using both SFA and PCA treatment approaches, recruitment of the left supramarginal gyrus and the right precuneus was associated with improved naming ability in a group of 8 PWA. Nardo, Holland, Leff, Price and Crinion (2017) also identified similar bilateral activation patterns in the language and domain-general networks which correlated with both immediate and long-term success in a group of 18 participants with chronic aphasia. In this research paper, the authors report significant improvements in the accuracy and speed of naming following a phonological cueing therapy. In the fMRI paradigm, the therapy task was mirrored as participants were required to name pictures with the aid of three types of phonemic cues (whole word, initial phoneme and final phoneme). Gains post-therapy were correlated with activation in the right anterior insula, the inferior frontal and dorsal anterior cingulate cortices and the left premotor cortex, indicating the importance of bilateral language and cognitive networks in anomia recovery.

Others have also provided evidence for recruitment of both the LH and the RH in recovery but suggest regions may be implicated based on lesion size and location. In a therapy study with three participants with chronic aphasia, Fridriksson, Morrow-Odom, Moser, Fridriksson and Baylis (2006) completed 10 four-hour word finding

treatment sessions using an error-reducing learning paradigm. Following therapy, two of the three participants improved but differential bilateral activation patterns were identified. For the participant with the larger lesion, gains post-therapy were associated with increased activity in the left temporal lobe and in the right inferior parietal lobe. For the other participant, who had the smallest lesion (primarily affecting Wernicke's area), naming improvements correlated with increased activity in the left parietal lobe and bilaterally in the frontal poles (the aPFC). As this frontal region has been associated with inhibition, the authors suggest that activation here reflected (not always successful) attempts to inhibit competing lexical targets during naming. Vitali and colleagues (2007) also identified differential patterns in relation to lesion characteristics, with one of their two participants, whose lesion partially spared Broca's area, showing treatment-induced activity in perilesional regions (par triangularis of the IFG) and in the right hippocampus, whereas the participant with complete damage to Broca's area showed increased post-therapy activation in the RH hemisphere homologue of the IFG.

Lastly, downregulation of neural activity in both the language and domain-general networks of the RH has also been associated with positive therapeutic outcomes, perhaps reflecting increased processing efficiency (Kiran & Thompson, 2019). In Abel, Weiller, Huber, Willmes and Specht's (2015) study, 14 participants with anomia as part of their aphasia took part in a word finding therapy programme utilising both semantic and phonological-based cueing hierarchies. At both pre-therapy and post-therapy, participants completed an fMRI paradigm which consisted of items trained in the therapy and untrained control items. Both therapy approaches significantly improved naming, though differential patterns were found in relation to impairment type and therapy success, as well changes in neural activation post-therapy.

Participants with a primarily semantic disorder benefitted similarly from both types of therapy and gains were correlated with decreases in activity in the left STG, caudate nucleus and paracentral lobule, as well as in the right rolandic operculum. Conversely, those with more of a phonological deficit responded better to phonological therapy and improvements in naming were associated with decreased activation in the LH in the STG and precuneus, in the RH in the caudate nucleus, and bilaterally in the thalami.

Furthermore, decreases in activation which correlate with naming success have been found in the RH in the IFG and insula following constraint-induced language therapy with 16 PWA (Richter, Miltner & Straube, 2008). Others, however, have found that both decreases and increases in neural activity change over time and may also depend on the therapy approach utilised and the intensity of the treatment programme (Menke et al, 2009; Marcotte et al, 2018). In Marcotte and colleagues' study, PCA was administered to two PWA, but one participant received intensive therapy and the other received PCA in a standard, non-intensive schedule. Although both participants improved, the one who received non-intensive therapy showed increased activation in the right precentral gyrus, the left IFG and bilaterally in the STG, and decreased activity in the MTG, the precentral gyrus and putamen of the LH, and in the postcentral gyrus of the RH. On the other hand, the participant who received intensive treatment showed greater activation in the right caudate nucleus and the left medial frontal gyrus, and decreased activity in the RH in the posterior cingulate gyrus, the precentral gyrus and the medial frontal gyrus. These differing results are reflective of the heterogeneity between PWA in treatment-induced neuroplasticity and highlight the difficulties with comparing treatment studies that employ different therapy approaches and intensity schedules.

1.5.3 Summary of neuroplasticity in aphasia

Overall, current research provides compelling evidence for the brain to adapt and reorganise following aphasic stroke, both spontaneously and in response to treatment. However, the inconsistent and contradictory findings are perplexing and increasingly challenging to unravel in order to identify the mechanisms underlying recovery. This is perhaps understandable considering the variability between studies in terms of treatment approach and imaging method in a population who are highly diverse at both the neural and the behavioural level (Doogan et al, 2018). Treatment-induced neuroplasticity studies in aphasia have also been dominated by case series, with only a few group studies, and it is unclear whether the results from one or two PWA can be used to make inferences across individuals (Li, Mukadam & Kiran, 2022). Future research employing multivariate approaches to understand the complex relationships between behaviour and the brain may provide a more comprehensive understanding of language recovery and neuroplasticity (Wilson & Hula, 2019). Ultimately, the path of language recovery in chronic aphasia remains unclear, but it does seem to be determined by the brain's ability to adapt and reorganise, developing compensatory mechanisms for disrupted language processes (Hartwigsen & Saur, 2019).

As a final important note, for the purposes of my thesis, the fMRI studies discussed above have focused on those using task-based paradigms and investigating relative changes in activation. However, aphasia is a network disorder, comprising of functionally and anatomically connected regions across the right and left hemispheres, and recovery of function following damage may be inadequately described by simply amount of neural activity (Kiran & Thompson, 2019; Seigel et al, 2018). Although the reported studies which utilised DTI examined structural changes

in white matter networks, crucial research also demonstrates treatment-induced changes in functional connectivity (Kiran & Thompson 2019). Be that as it may, further commentary on these connectivity-based analyses is outside the scope of the current study.

2 Overview of thesis

Each of the three results chapters in this thesis relate to the repeated measures Phase II clinical trial of the iTalkBetter word retrieval therapy app (Figure 2.1). Twenty-seven participants completed the study and, over the course of the six-week treatment period, they aimed to complete 60 hours of therapy. The study consisted of five behavioural testing time points (T1-T5), each two to 12 weeks apart. T1 was the baseline time point; the interval between T2-T3 formed the pre-therapy, or no therapy, block; T3-T4 formed the therapy block; and T4-T5 was the follow-up period to investigate whether any improvements were maintained three months later without the iTalkBetter therapy. Seventeen participants who were medically suitable and available for scanning (dependent on the COVID-19 situation), completed structural and functional (task-based and resting state) magnetic resonance imaging (MRI) scanning before (T2 and T3) and after (T4) therapy.

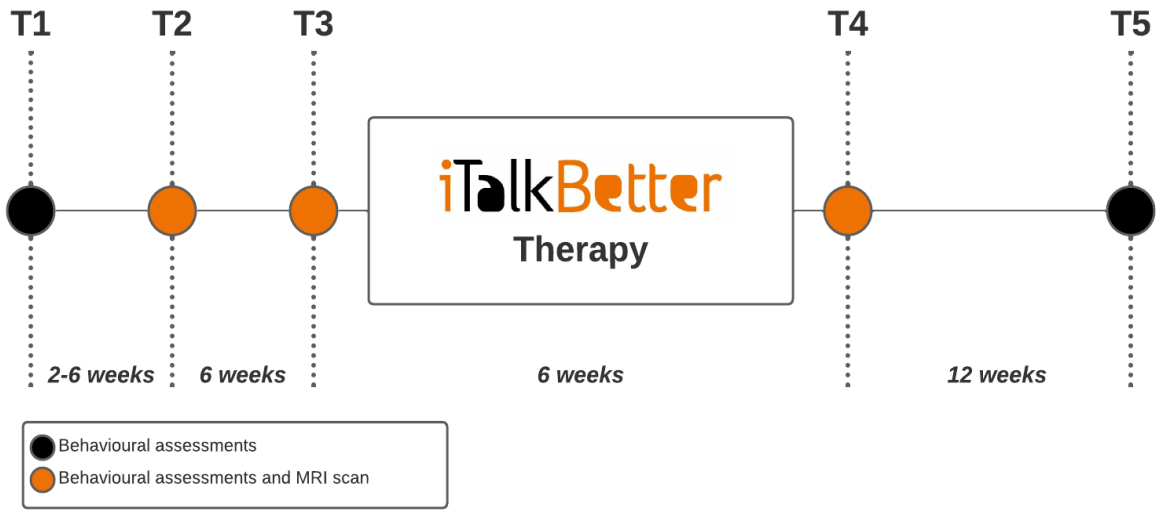


Figure 2.1 iTalkBetter study design.

3 Main aims and hypotheses

The results in this thesis are presented in three chapters. The aims and hypotheses for each of these chapters are detailed below.

3.1 Chapter 1: Investigating the behavioural response to iTalkBetter therapy in persons with chronic aphasia.

In light of the above research evidence demonstrating the efficacy of high-dose, self-led, computerised therapy, the current study will examine the effectiveness of one such therapy in improving word retrieval deficits in a group of participants with chronic aphasia. This therapy, called iTalkBetter, utilises a phonological, error-reducing cueing paradigm and incorporates a variety of highly frequent words (both abstract and concrete) due to the research indicating the benefits of these components within anomia intervention.

Aim 1

My primary aim is to investigate whether iTalkBetter can improve single word retrieval in persons with chronic aphasia.

Aim 2

My second aim is to examine if improvements in word retrieval in a single word naming test generalise to improvements in word retrieval in a spoken picture description task.

Aim 3

My final aim is to explore if baseline factors can explain response to therapy.

Hypotheses:

(1) At the group level, participants will significantly improve their retrieval of trained lexical items following iTalkBetter therapy and this improvement will not generalise to (psycholinguistically matched) untrained items.

Previous studies have shown that single word anomia therapy is effective when given at high dose, however, there is a large evidence-base demonstrating that improvements do not often generalise to untrained words. I predicted this would also be the case in the current study due to the impairment-based therapy approach employed.

(2) At the group level, participants will significantly improve their ability to retrieve both trained concrete and trained abstract lexical items.

There is limited evidence available that traditional therapy techniques may also be suitable in targeting more abstract words (Renvall & Nickels, 2019). The present study builds on this evidence by investigating changes in both abstract and concrete words following a traditional intervention approach with a larger cohort of participants. It was expected that improvements would be significant regardless of word type.

(3) Following iTalkBetter therapy, participants will significantly improve their retrieval of trained lexical items in a spoken picture description task.

Theoretically, improvements at the single word level should generalise to improvements in connected speech, though the support for this in the aphasia literature is sparse. It has been suggested that generalisation across communication contexts is more likely if the task is similar to that targeted in therapy (Conroy, Sage and Lambon Ralph, 2009c). Therefore, I predicted participants would generalise their

learning from single word picture naming to spoken picture description (SPD) due to the similarities between these tasks (both are static pictures which require a verbal response).

(4) A combination of demographic and baseline behavioural factors will explain individual response to the iTalkBetter therapy.

Previous studies have shown the variability between PWA in their response to therapy and how baseline factors may explain some of this variance. However, which baseline factors contribute to this variability remains unclear. Therefore, I expected that a combination of factors would explain differences in individual response to iTalkBetter.

3.2 Chapter 2: Investigating structural brain adaptation in response to iTalkBetter therapy.

In healthy individuals, longitudinal structural MRI imaging has been utilised to demonstrate subtle changes in regional brain tissue in response to training. However, to date, there is limited research that has investigated this in PWA following therapy intervention. Due to this, Chapter 2 explores therapy-induced structural neuroplasticity in a subset of participants who took part in the iTalkBetter study.

Aim

My aim in this Chapter is to investigate whether iTalkBetter therapy induces regional changes in brain tissue in persons with chronic aphasia.

Hypotheses

(1) iTalkBetter therapy will induce increases in volume in the grey and/ or white matter of the language and/ or cognitive networks.

There is a paucity of studies which have examined therapy-induced neural reorganisation using structural imaging techniques. The ones that have identified changes in both the language network in the LH and in language homologue regions in the RH. Furthermore, in functional imaging, changes in relative activation have also been found in the cognitive networks following intervention. For these reasons, I hypothesised that, following six weeks of iTalkBetter therapy, tissue changes would be observed in either the left or the right hemisphere in the language or cognitive networks.

(2) Increases in volume in the grey and/ or white matter will correlate with percentage change on trained items from pre- to post-therapy.

Changes in tissue volume following aphasia therapy were found to correlate with changes in language function in three of the four studies discussed in the literature review which utilised structural imaging (Wan et al, 2014; Chang et al, 2021; Fleming et al, 2020). Due to this, I predicted that any therapy-driven changes in the structural architecture of the brain would be associated with treatment outcomes in the iTalkBetter study.

(3) Dose of therapy (number of hours completed) will correlate with increases in the volume of grey and/ or white matter.

In experience-driven neuroplasticity with healthy individuals, changes in brain structure have been associated with the amount of training (Golestani, Price and Scott, 2011). As iTalkBetter is a repetitive, high-dose treatment, I expected that

changes in the right and left hemispheres would be related to the amount of therapy participants achieved in the study.

3.3 Chapter 3: Investigating changes in task-related brain activity in response to iTalkBetter therapy.

In the aphasia literature, a number of studies have explored therapy-driven changes in brain activation in persons with chronic aphasia. However, differential findings have been reported, with both regions in the left and right hemispheres increasing and/or decreasing in activity following therapy. Therefore, this Chapter will examine functional brain activation from pre- to post-therapy to attempt to elucidate, at the group-level, regions of relative change in response to iTalkBetter.

Aim

In this Chapter I will be investigating whether iTalkBetter therapy induces changes in the language and/ or cognitive networks, identified using fMRI scanning during speech production tasks, in persons with chronic aphasia.

Hypotheses

(1) iTalkBetter therapy will induce changes in functional activation in the language and/ or cognitive networks, identified using speech production tasks (bilateral frontal and/ or temporal lobes).

Although there are many more studies which have examined therapy-induced neuroplasticity using functional imaging methods than structural imaging techniques, there is large variability in where activity changes are found. Due to this, I predicted iTalkBetter therapy would induce changes in brain activation, however, I kept this

hypothesis broad to include both the right and left hemispheres in the frontal and/ or temporal lobes.

(2) Changes in task-related activation will correlate with percentage change on trained items on the WRT following therapy.

Many studies have also correlated changes in task-related activation with behavioural improvements following therapy. Therefore, I expected this would also be the case in the iTalkBetter trial.

(3) Dose of therapy (number of hours completed) will correlate with changes in task-related activation.

As in the structural imaging, due to the repetitive and high-dose nature of the iTalkBetter therapy task, I predicted that changes in brain activity following intervention would be related to the amount of therapy participants completed.

4 Methods

4.1 Design

The iTalkBetter clinical study was a small-scale, repeated measures, control trial in which participants completed 6 weeks of therapy with the iTalkBetter app, aiming to complete 60 hours of therapy. The study consisted of five testing time points, each two to 12 weeks apart (Figure 4.1). T1 was the baseline time point; the interval between T2-T3 formed the pre-therapy, or no therapy, block; T3-T4 formed the therapy block; and T4-T5 was the follow-up period to investigate whether any improvements were maintained three months later without the iTalkBetter therapy.

At all time-points, participants completed the primary outcome measure: a novel naming assessment, the Word Retrieval Test (WRT). At T2, T3 and T4, a novel Spoken Picture Description task (SPD), and structural and functional MRI scans, for those who were suitable and available for scanning, were acquired. Both pre-therapy (T3) and post-therapy (T4), semi-structured interviews were completed, and at T1, T2 and T3, baseline language and cognitive assessments were administered. Due to the COVID-19 pandemic, seven of the participants' testing sessions were completed entirely remotely via the video calling platform, Zoom (https://zoom.us/). For the remaining 20 participants, testing sessions were either completed via Zoom, or were completed in-person at the Institute of Cognitive Neuroscience, University College London (UCL), the Wellcome Centre for Human Neuroimaging, UCL, or in participants' homes.

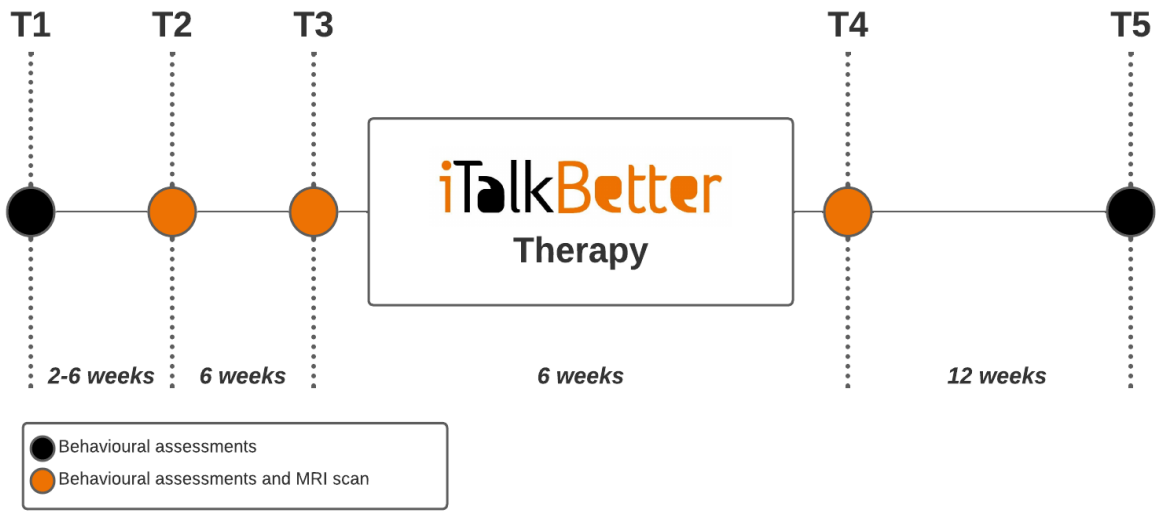


Figure 4.1 iTalkBetter study design.

4.2 Participants

31 PWA were recruited for the study. Four participants' data were excluded from analyses: two participants withdrew following T2 due to the change in the COVID-19 situation; one participant withdrew following T3 due to illness; and one participant unfortunately passed away following T4. Therefore, in this thesis, the data presented are for 27 participants (Table 4.1). Recruitment was from the Predicting Language Outcome and Recovery After Stroke (PLORAS) database (Wellcome Trust Centre for Human Neuroimaging, UCL (Price, Seghier & Leff, 2010)) and from a local outpatient clinic.

4.3 Inclusion and exclusion criteria

The inclusion criteria were as follows: (i) chronic post-stroke aphasia (at least six months post-stroke); (ii) adults, aged 18 or over; (iii) English as a dominant language; (iiii) anomia in the absence of a severe speech output deficit as evidenced by: (a) impaired word retrieval on the 'Object naming' subtest of the Comprehensive Aphasia Test (CAT) (Swinburn, Howard & Porter 2004) (cut-off <38); and (b)

relatively intact single word repetition on the 'Repetition' subtest of the CAT (scores >12). Exclusion criteria were: (i) diagnosis of developmental language disorders; (ii) major co-existing neurological or psychiatric disorders; (iii) unable to give informed consent. Written informed consent to take part in the study was obtained from each participant.

ID	Gender	Age	Handed-ness	CAT-naming	CAT-repetition	Time since stroke (months)	Type of stroke	Lesion volume (cm^3)
P1	M	65	R	29	21	109	Haemorrhagic	158
P2	M	58	R	16	16	90	Ischaemic	168
P3	M	70	R	13	30	91	Ischaemic	244
P4	F	62	R	29	24	22	Ischaemic	52
P5	M	64	L	3	23	14	Unknown	232
P6	M	59	R	34	29	100	Ischaemic	283
P7	M	57	R	28	20	132	Ischaemic	-
P8	F	82	R	27	23	38	Ischaemic	156
P9	M	69	L	32	19	136	Ischaemic	-
P10	M	56	R	26	24	156	Ischaemic	-
P11	M	62	R	24	28	149	Ischaemic	-
P12	M	64	R	0	26	62	Ischaemic	-
P13	M	64	R	17	28	74	Ischaemic	-
P14	F	31	L	33	28	60	Haemorrhagic	-
P15	M	68	R	30	19	184	Ischaemic	-
P16	M	64	R	28	22	79	Ischaemic	158
P17	F	56	R	25	24	33	Ischaemic	352
P18	M	73	R	35	30	43	Ischaemic	109
P19	F	61	R	1	13	30	Ischaemic	150
P20	M	67	R	26	22	320	Ischaemic	376
P21	M	60	R	32	30	43	Ischaemic	166
P22	M	70	L	29	27	106	Unknown	264
P23	M	64	R	16	19	31	Ischaemic	130
P24	F	54	R	29	27	37	Ischaemic	32
P25	F	55	R	30	32	19	Ischaemic	191
P26	M	45	R	35	32	7	Ischaemic	76
P27	M	64	R	6	18	84	Ischaemic	386
Mean		62		23	24	83		194
SD		9		4	5	67		103
Max				48	32			
Cut-off				38	12			

Table 4.1 Participant demographics.

4.4 Ethics

Ethical approval for the iTalkBetter study was granted by the National Research Ethics Service Committee East of England, Cambridge (18/EE/228).

4.5 iTalkBetter therapy task

iTalkBetter is a digital word retrieval therapy app, which was completed independently at home on a computer tablet. It consists of single word picture naming that works via mass practice using error reducing learning (vanishing cues). iTalkBetter also utilises gamification through the use of an 'outer space' theme. Gamification was used to reduce the boredom and fatigue effects often associated with repetitive and intensive tasks, thereby decreasing participant drop-out rates (Brady, Kelly, Godwin, Enderby & Campbell, 2016; Dignam, Rodriguez & Copland, 2016).

4.5.1 Error-reducing learning

Picture naming paired with error-reducing, vanishing phonological cues has been shown to facilitate word retrieval following therapy with long-term benefits (Conroy, Sage & Lambon Ralph, 2009b; Nardo, Holland, Leff, Price & Crinion, 2017). For this reason, vanishing cues were incorporated into the iTalkBetter therapy task (Figure 4.2). In the task, when participants first saw a picture, they heard the name of the picture (full word cue (FC)). They then had to repeat the picture's name. If the correct name was produced, the participant saw the image again but heard only the initial phoneme (initial phoneme cue (IP)). They again had to say the whole name. If the name was said correctly, the participant saw the picture for a third time but received no auditory cue (no cue (NC)) and once more, they had to say the whole name of the picture.

If at FC a participant said the incorrect name, they were given two further instances to name the picture at FC before the therapy moved on to the next word. If an incorrect name was produced at either IP or NC, the picture was presented again at the previous cue level. In the next therapy cycle, the iTalkBetter therapy presented the picture at the cue level the participant achieved in the previous therapy cycle (see section 4.5.4 for details on the therapy cycle). For example, if the name was produced correctly at NC in the previous cycle, the picture was first presented at NC; if correct in the previous cycle at IP, the picture was first presented at IP; and if the participant was unable to move beyond FC, the picture was presented at FC.

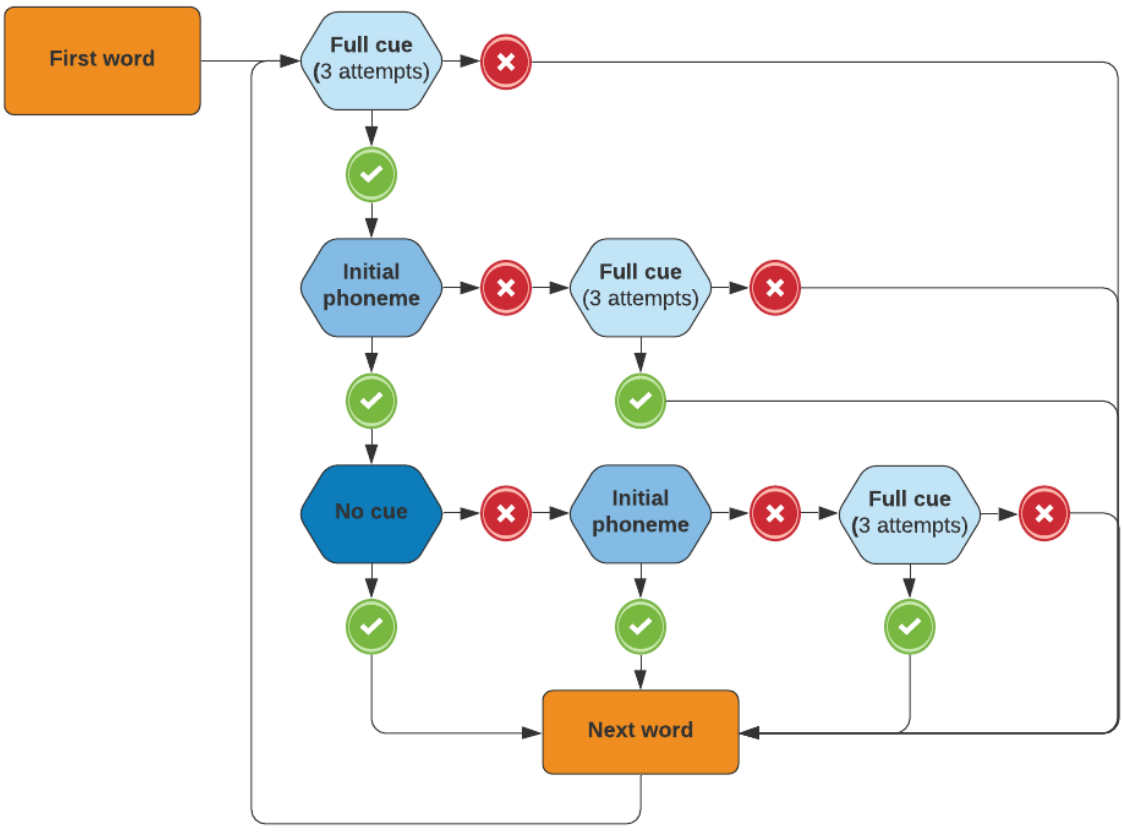


Figure 4.2 Error-reducing learning.

4.5.2 Speech recogniser

A novel automatic speech recogniser was developed by a member of the Neurotherapeutics group at UCL and was incorporated into the iTalkBetter therapy (Barbera et al, 2021). This naming utterance verification system (NUVA) utilised a deep learning element that classified whether the correct or incorrect name was produced in real time. This allowed the app to determine the next cue level and also enabled the provision of immediate feedback on a trial-by-trial basis, thereby improving word retrieval outcomes (McKissock & Ward, 2007). If a user produced the correct name, 'Well done!' appeared on the screen and 'Kenny', the spaceman, hoovered up the picture into his 'space backpack'. If the incorrect name was produced, the picture floated off into 'outer space' (Figure 4.3).

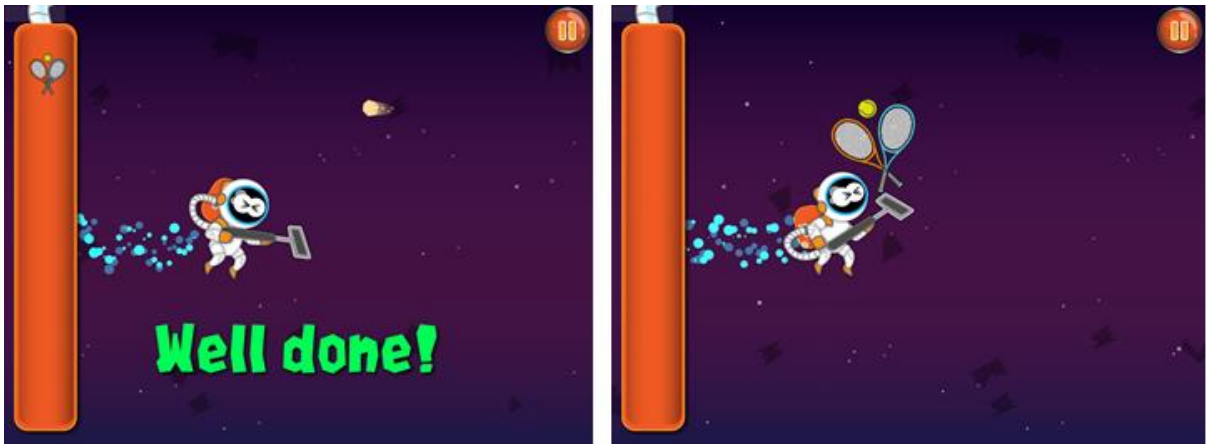


Figure 4.3 Therapy feedback. Left: feedback for correct response. Right: feedback for incorrect response.

4.5.3 Therapy content

The therapy content was a list of target words ('lexical items') composed of all major word categories, including nouns, verbs, adjectives, prepositions and pronouns. A novel feature of iTalkBetter was that both abstract words (words that refer to intangible concepts (often words with low imageability), for example, 'confidence')

and concrete words (words that portray tangible concepts (often highly imageable words), for example, ‘cat’) were included in the therapy. Previous picture naming interventions have predominantly included only concrete lexical items. Abstract words, however, are as prevalent in the English language as concrete words and are highly integral to the ability to communicate and convey thoughts, feelings and emotions (Binney, Zuckerman & Reilly, 2016).

4.5.3.1 Initial word list

Generation of the initial word list, which was given to the first seven participants in the study, was completed using the British English SUBTLEX-UK database (van Heuven, Mandera, Keuleers & Brysbaert, 2014). The 2000 most frequent words in the English language were identified. These 2000 words were put into randomised lists which were distributed to members of the Neurotherapeutics group at UCL (n=7) and PWA (n=4). Each person chose 1000 words which they thought would be the most useful to learn and the top 900 were selected for the therapy content. The remaining 100 words were chosen from a previous study which asked 100 PWA what words they would want to learn in therapy (Palmer, Chater & Hughes, 2017). It was important to choose words that people had indicated they would want to learn as impairment-based therapies (such as iTalkBetter) often show item-specific learning effects (Middleton et al, 2016).

4.5.3.2 Secondary word list

For the following 22 participants who took part in the iTalkBetter study, the word list was greatly reduced due to the time it took to work through the entire therapy content. Over the course of the six weeks of iTalkBetter therapy, the first seven participants were exposed to each lexical item an average of six times. Due to this

limited exposure, the therapy content was condensed to 220 items. For each participant, this consisted of 110 items from the primary outcome measure (see section 4.6.1), the 100 items from Palmer, Chater and Hughes' study, and 10 items which were randomly selected from the initial word list.

4.5.4 Therapy cycle

At the start of iTalkBetter, the order of exposure to the therapy content was randomised and separated into blocks ('planets') that consisted of 20 lexical items. Within a planet, participants practiced naming each picture separately, moving through the cue levels as described in section 4.5.1. Once all lexical items within a planet were complete, participants moved on to the next planet to practice the next set of 20 items (Figure 4.4). This process continued until all the items in the therapy content were presented (one therapy cycle). The items were then re-randomised and the process began again.

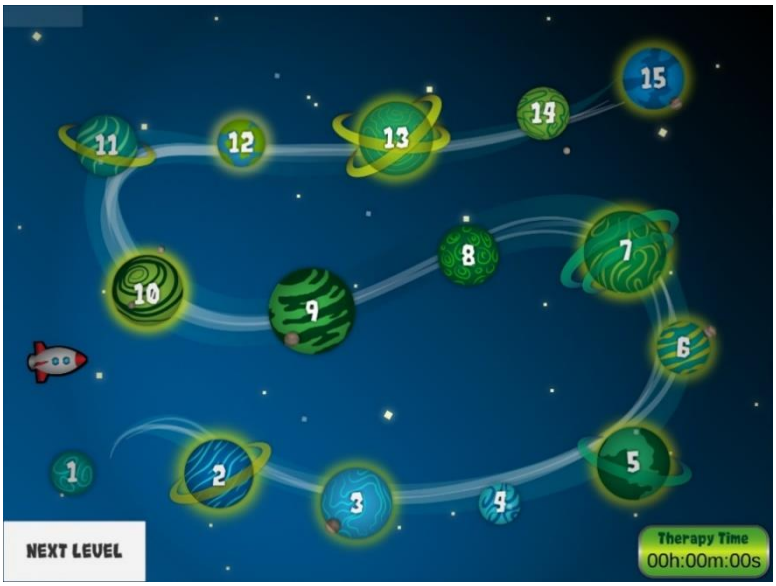


Figure 4.4 Therapy planet progression.

4.6 Behavioural assessments

4.6.1 Word Retrieval Test

The Word Retrieval Test (WRT) is a novel assessment developed for the iTalkBetter therapy and was used as the primary outcome measure for the study. It is a single word picture naming assessment that was completed at all time-points (T1-T5). The WRT was created with two main requirements: inclusion of a large number of lexical items to be sensitive in capturing change, and inclusion of trained and untrained items to assess whether treatment effects were item-specific or if they generalised to untrained items. The WRT was shown to have high concurrent validity with the 'Object naming' subtest of the CAT, r=0.733, p=<0.001, suggesting the test is a valid measure of single word naming.

4.6.1.1 Content

The test consisted of 220 lexical items (nouns, adjectives, verbs and pronouns) which were chosen from the initial therapy content to be representative of the entire initial corpus based on the following variables: word class, syllable length, concreteness (Brysbaert, Warriner & Kuperman, 2013), frequency (van Heuven et al, 2014) and age of acquisition (Kuperman, Stadthagen-Gonzalez & Brysbaert, 2012). The pictures in the WRT, however, were different to those in the therapy to avoid participants' rote learning the association between the spoken word and the picture (identity priming) (Schacter, Dobbins & Schnyer, 2004).

4.6.1.2 Trained and untrained words

The 220 lexical items in the WRT were matched to one another according to the above psycholinguistic factors to produce 110 matched word pairs. Baseline

performance on the WRT at T1 and T2 was then used to assign words within a pair to either trained or untrained word lists by balancing performance between pairs (Table 4.2). The 110 words which were selected to be untrained were removed from the therapy content and the 110 trained words were kept in the therapy content. This allocation method created unique lists of trained and untrained items for each participant which were matched for both key psycholinguistic variables and baseline performance.

Pair 1	**T1**	**T2**	**Pair 2**	**T1**	**T2**
bread	1		teeth	1	1
girl	1		train	1	1
wheel	1	1	juice		1
knife	1	1	snow		1
chicken	1	1	orange		1
upstairs	1	1	downstairs		1
soup	1	1	flag	1	
gift	1	1	hole	1	
answer			return		
shops	1	1	nurse	1	1
sight		1	spring		
joke			mind		1
TOTAL (TR)				(T1) **7**	(T2) **8**
TOTAL (UN)				(T1) **7**	(T2) **8**

Table 4.2 Example allocation of trained and untrained items. Each row is a word pair. A '1' in columns T1 and T2 denotes correct production of the word. Green: words allocated to the trained items list; red: words allocated to the untrained items list.

4.6.1.3 WRT components

Additionally, as the therapy content included abstract words, the WRT had to be devised in such a way that it would be reliable in evaluating improvements in the retrieval of words with low imageability. Although the pictures were designed to depict the target words as clearly as possible, for abstract words it was still difficult to infer the meaning of a picture without first knowing what it represented (Figure 4.5). Therefore, in a simple picture naming test, any improvements in the production of

abstract words following therapy could be due to learning the representational meaning of a picture, rather than to an improvement in word retrieval.

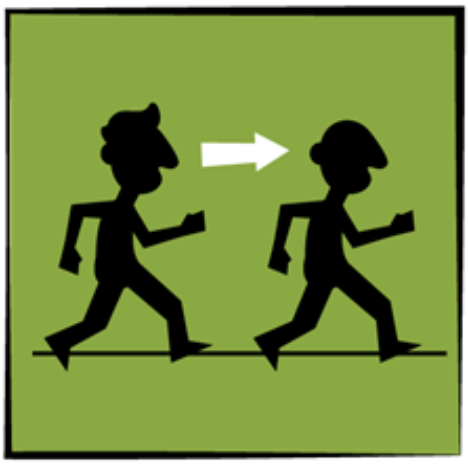


Figure 4.5 Abstract word picture example: 'following'.

To guarantee improvements in naming following therapy were due to improvements in word retrieval and not to improvements in understanding the meaning of a picture, two distinct components were incorporated into the WRT: a 'fly by' in which participants heard the names of 20 pictures; and a 'naming test' in which participants saw the same 20 pictures again and had to free name the pictures (see Figure 4.6).

In the fly by, each of the 20 pictures assigned to the block were shown one at a time for six seconds whilst the audio recording which corresponded to the picture was played. During this time, 'Kenny', the spaceman, was sat on the floor with his hand to his ear. Once all 20 pictures had been shown, the beginning of the naming test was signaled by the spaceman jumping up and running. In the naming test, the 20 pictures from the fly by were randomised and shown again for six seconds each. Participants were required to say the name of the picture (which they had been previously given in the fly by segment) before the test automatically moved on to the next picture.

To ensure consistency in exposure to items, all 220 words (both trained and untrained; abstract and concrete) were included in both components. In the test, there were 11 blocks which consisted of a 'fly by' and a 'naming test', and at the end of each block, participants were able to have a break.

Figure 4.6 The WRT: Left: Listening to the picture names (the 'fly by'). Right: Free naming the pictures (the 'naming test').

4.6.1.4 The fly by: Amazon Mechanical Turk

The block length of 20 words was selected following the completion of an online experiment with healthy aged matched controls. 38 participants (13 male), aged 57 to 72 years (M = 61.6, SD = 3.73) who spoke English as their dominant language took part in the experiment via the platform Amazon Mechanical Turk (https://www.mturk.com/). In the task, participants were presented with blocks ('fly by' followed by 'naming test') of word lists of differing lengths (10, 15, 20, 25 and 30 words). Each block length was presented twice in a random order and all blocks contained randomised lists of different words.

For the WRT, the block length with an overall accuracy rate of 80-90% was determined to be the 'sweet spot' to balance the memory component of the task (both working memory and episodic memory) which becomes harder with increasing

block length, with priming effects, which dominate shorter block lengths. Blocks of both 20 and 25 words were within this bracket (86.62% and 84.79%, respectively) (see Figure 4.7). There was no statistically significant difference in the accuracy between these two blocks, $t(37) = 0.64$, $p = 0.52$, so, to reduce the working memory load and as the WRT consists of 220 words and 20 is a factor of 220, this block length was chosen.

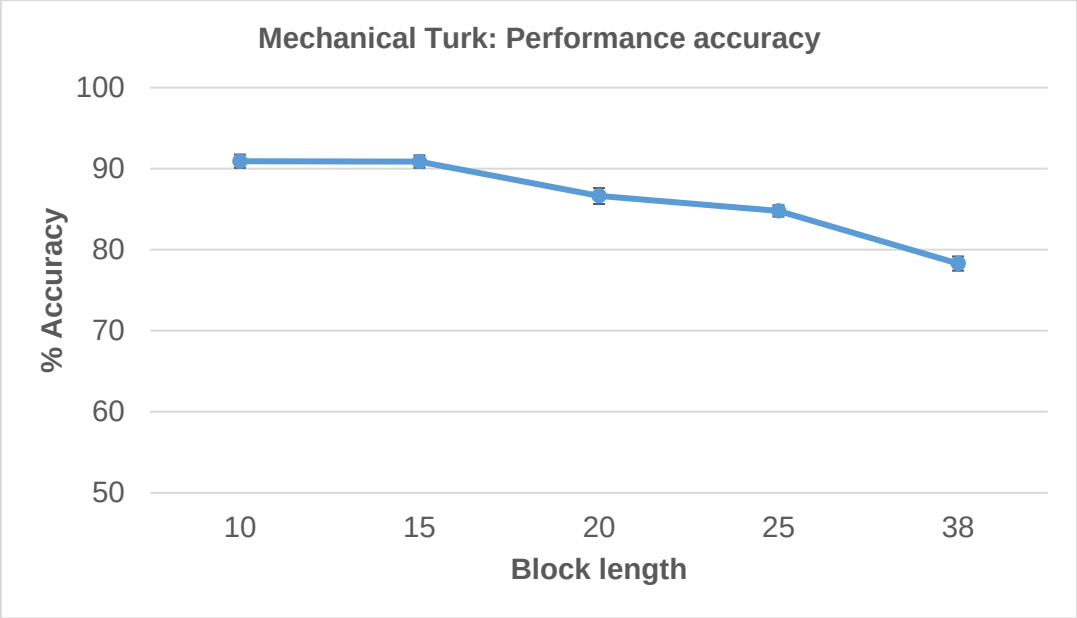


Figure 4.7 Mechanical Turk: Performance accuracy. Blue line represents the mean % accuracy score for each block. Error bars are within-subject standard error of the mean.

4.6.1.5 Presentation and scoring

The WRT was incorporated into the iTalkBetter app and at all testing time points the test was completed on a computer tablet via the app. The test audio recorded participants' responses and, following the test, responses were scored by hand (not the automated speech recogniser, which was used for therapy progression only). This included a full written transcription of each response which was then error coded into one of the following categories: correct, self-correct, late, morphological error, synonym, semantic error, circumlocution, phonemic error, visual error, mixed

error, unrelated error, perseveration, partial phonemic and no response (see Appendix A for error type descriptions and examples).

4.6.2 Spoken Picture Description

A novel Spoken Picture Description (SPD) task was used as a secondary outcome measure and was completed at T2, T3 and T4 (Figure 4.8). This assessment was created as a way to investigate whether any gains seen at the single word level (in the WRT) generalised to another, more ecologically valid, speech production task.

4.6.2.1 Trained and untrained words

The SPD consisted of two pictures of composite scenes that depicted a selection of trained and untrained words from the WRT to enable the direct comparison of naming performance between the two assessments. Items from the WRT were selected only if both words in the matched pairs (described in section 4.6.1.2) could be included (see Appendix B for word lists corresponding to each scene). As one word in a pair was always trained and the other was always untrained, this ensured an equal number of trained and untrained words from the WRT were incorporated into the SPD for each participant. The scenes also included trained items from the core therapy content that were not tested in the WRT, and items that were not trained or tested in the iTalkBetter study.

Figure 4.8 Spoken Picture Description. Top: Scene 1. Bottom: Scene 2.

4.6.2.2 Presentation and scoring

Each SPD was presented digitally as a PDF. Participants were given two minutes to describe what was happening in each picture and were instructed to try to use full sentences. Responses were audio recorded and then transcribed following the assessment sessions. The transcriptions were analysed by counting the number of appropriate: WRT words (trained and untrained words); core therapy words (trained in the therapy but not tested in the WRT); and untrained and untested words (words not trained in the therapy or tested in the WRT). The number of unique information carrying words were counted for each word type, any repetitions of words were not included in the analysis.

4.6.3 Dynamic Visual Analogue Mood Scales

The Dynamic Visual Analogue Mood Scales (DVAMS) (http://dvams.com) are a non-verbal assessment of mood, designed for people with post-stroke aphasia, which were completed pre-therapy (T3) and post-therapy (T4) (Barrows & Thomas, 2017). It consists of the following seven scales: (1) Miserable–Satisfied; (2) Sad–Happy; (3) Distressed–Peaceful; (4) Bored–Excited; (5) Afraid–Calm; (6) Angry–Peaceful; and (7) Sleepy–Alert. In the assessment, participants reported their mood by modifying facial expressions using a slider on a computer or tablet and this corresponded to a score, ranging from zero to 100. A score of zero marked the negative end of the scale, whereas a score of 100 marked the positive end of the scale. These scores were then totalled and averaged across the seven scales to give a final score of mood.

4.6.4 Semi-structured interviews

Semi-structured interviews (patient and carer reported outcomes) were also completed at pre-therapy (T3) and post-therapy (T4). These consisted of nine open-ended questions relating to the study, such as why the participants volunteered for the study and why they think research is important; what they were hoping to achieve/ what they did achieve in the study; and how this may impact/ has impacted their lives and their carers' or partners' lives. At the beginning of each interview, a brief introduction was given, followed by the nine questions, and, at the end, the researcher gave a summary of what was discussed and asked if there was anything the participant or carer would like to add. All interviews were audio recorded and are being analysed as part of an ongoing qualitative research project. The interviews will not be discussed further in my thesis.

4.6.5 Baseline tests

Further linguistic and cognitive assessments that were conducted are displayed in Table 4.3. These tests were administered once pre-therapy at either T1, T2 or T3. Three of the tests, and certain subtests of the CAT, could not be completed remotely, therefore, the participants who were not able to attend any in-person testing sessions due to COVID-19 lockdowns did not complete these assessments (n=7). These are denoted by an asterisk.

Assessment	Abbreviation	Domain
Linguistic		
Comprehensive Aphasia Test	CAT	Cognition*, comprehension, repetition, naming, reading and writing*
Psycholinguistic Assessments of Language Processing in Aphasia (Subtest 9)	PALPA	Non-word repetition
Cognitive		
Sustained Attention to Response Task*	SART	Sustained attention and inhibition
Cattell Culture Fair Intelligence Test Scale II: Tests 1 & 2*	CFIT	Non-verbal fluid intelligence
4-Way Weigl*	4-WW	Executive function

Table 4.3 Baseline assessments.

4.6.5.1 Comprehensive Aphasia Test

The CAT is a standardised assessment of multiple language domains that is commonly used in clinical and research settings to provide an overview of a person's speech and language profile. In the iTalkBetter study, an electronic version was used and the full CAT, excluding the disability questionnaire, was administered to those who were able to attend in-person testing sessions on a Windows laptop computer. For those who completed the full study remotely, the test was presented over Zoom. For these participants, a partner or carer was present to enable the scoring of the comprehension subtests, and the writing and cognition (apart from Verbal Fluency) subtests were not completed.

4.6.5.2 PALPA (non-word repetition)

The PALPA is another standardised battery of language tests which assesses language processing and all participants in the study completed the non-word repetition subtest (subtest 9) (Kay, Lesser & Coltheart, 1992). In this subtest, participants listened to 80 pre-recorded non-words, one at a time, and were asked to

repeat what they heard. They were allowed one repetition of the word if required and all responses were audio recorded and transcribed offline. A correct response scored one and an incorrect response scored zero, giving a final score out of 80.

4.6.5.3 Sustained Attention to Response Task

The SART is a test of sustained attention and inhibition and a non-verbal version was used in the iTalkBetter study (Manly, Davison, Heutink, Galloway, & Robertson, 2000). It is a Go/No-Go task, presented on a Windows laptop, involving 216 trials of pictures of two different men. Whenever participants saw the picture of one of the men, they were required to press a button as quickly as they could (Go trial; N=192), but when they saw the picture of the other man, they were required to withhold their response (No-Go trial; N=24). The Go/No-Go trials were presented in a pseudorandomised order and, as No-Go trials were unpredictable and uncommon, participants had to sustain their attention throughout the assessment. The total error score for this test (SART-errors) was the number of false negative (incorrect omissions on Go trials) and false positive (incorrect hits on No-Go trials) responses.

4.6.5.4 Cattell Culture Fair Test Scale II: Tests 1 and 2

The CFIT is a measure of fluid intelligence and was created to be a non-verbal and culturally neutral test, so that it can be used with those who do not have English as their dominant language, or have language processing impairments (Cattell & Cattell, 1949). Subtests 1 and 2 from scale II were administered in the present study. In subtest 1, participants were presented with four black and white drawings and were required to select a fifth picture to complete a sequence (N=12). In subtest 2, participants were presented with five black and white drawings and had to select the odd one out (N=14). Both subtests were timed (subtest 1: three minutes; subtest 2: 4

minutes) and the final scores were calculated as a composite score from the two subtests.

4.6.5.5 4-Way Weigl

The 4-WW is a token sorting test which assesses executive functions such as set shifting, problem solving, perseveration and cognitive flexibility (Beglinger, Unverzagt, Beristain, & Kareken, 2008). In the test, 16 plastic tokens were randomly arranged and participants were instructed to sort the tokens into groups so that all tokens within a group were the same in one way. There were four possible sorts to complete: by colour, shape, texture or by the symbol on the top of the token. If a participant was unable to complete a sort, the researcher provided assistance by completing one group in a sort (for example, putting all of the yellow tokens together if sorting by colour) (Step down A). If a participant was still unable to complete a sort, further assistance was given in the form of verbal confirmation of the type of sort (for example, 'by colour') (Step down B). Participants were awarded three points if the sort was completed unassisted; two points if the sort was completed after Step down A; and one point if completed after Step down B (maximum score: 12). The number of failures to complete a sort after both step downs were also counted, as were the number of perseverative errors (repeating a previously completed sort).

4.7 Chapter 1: Behavioural data analysis

All behavioural data analyses were completed in the Statistical Software Package for the Social Sciences 25 (SPSS).

4.7.1 Response to therapy

The primary research question was whether there was a significant improvement in single word retrieval, as measured by the WRT, following six weeks of iTalkBetter therapy compared to six weeks of no therapy. To investigate this question, change scores from T2-T3 (pre-therapy) and T3-T4 (therapy) on the WRT were analysed using repeated measures analysis of variances (ANOVAs) and the details of the models are reported in Chapter 1.

4.7.1.1 Baseline stability

To examine baseline stability on the WRT, paired sample t-tests were completed using performance scores at T1, T2 and T3.

4.7.1.2 Maintenance effects

The maintenance of any treatment effects was also explored using a paired sample t-test for performance at T4 (immediately post-therapy) and T5 (12 weeks following therapy).

4.7.2 Generalisation of response to therapy

To investigate whether any improvements seen at the single word level on the WRT generalised to the SPD task, change scores from T2-T3 (pre-therapy) and T3-T4 (therapy) on the SPD were also analysed using a repeated measures ANOVA, which is described in Chapter 1.

4.7.3 Explaining response to therapy

To explore the associations between key demographic and behavioural variables, and response to therapy, simple correlations were employed.

As multiple baseline tests were completed before the start of therapy, explanatory modelling was also used to examine if a combination of baseline variables (both behavioural and demographic) could explain response to therapy. Automatic Linear Modelling (ALM) was chosen for this as it performs linear regression analysis using multiple predictor variables. ALM is advantageous over traditional linear regression techniques as it generates an optimal model by considering all possible combinations of potential predictors for the dependent variable (which in this case is response to therapy) (Field, 2013). The ALM procedure is automatically optimised for data preparation and with two steps: 1) data preparation; 2) variable selection.

In data preparation, any missing values are replaced, categorical variables are transformed and the Cook's distance value is determined. The Cook's distance value is calculated for any outliers which are three or more standard deviations away from the mean and is an estimation of the impact of the outlier(s) on the model. If any values are close to one, this indicates a greater impact on the model and the exclusion or inclusion of these cases should be examined (Field, 2013). In variable selection, ALM provides two possible techniques: stepwise selection and best subtests, which is used in this thesis. Best subtests considers all possible models and is recommended when the number of predictor variables is less than 20 (Yang, 2013).

The adequacy of the final model is estimated by the following parameters: adjusted R^2 (a measure of how much variance is explained by the model); Akaike's information criterion (AIC) (a model comparison measure which penalises for more complex models: the model with the lowest AIC is the best model); average squared error (prevents model overfitting); and F-test (a model comparison which assesses which model best fits the data).

4.8 Magnetic Resonance Imaging

MRI is a non-invasive technique which generates high quality images of the fine details of internal anatomy. In the iTalkBetter study, participants who were able and available for MRI scanning completed structural and functional scans at T2, T3 and T4. In the following sections, a brief overview of the basic principles of MRI and the techniques employed in this study for acquiring and analysing structural and task-based functional MRI are provided.

4.8.1 The MRI scanner

The MRI scanner consists of four crucial components: a superconducting electromagnet; radio-frequency (RF) transmitter and receiver coils; three sets of gradient coils; and shimming coils (Figure 4.9).

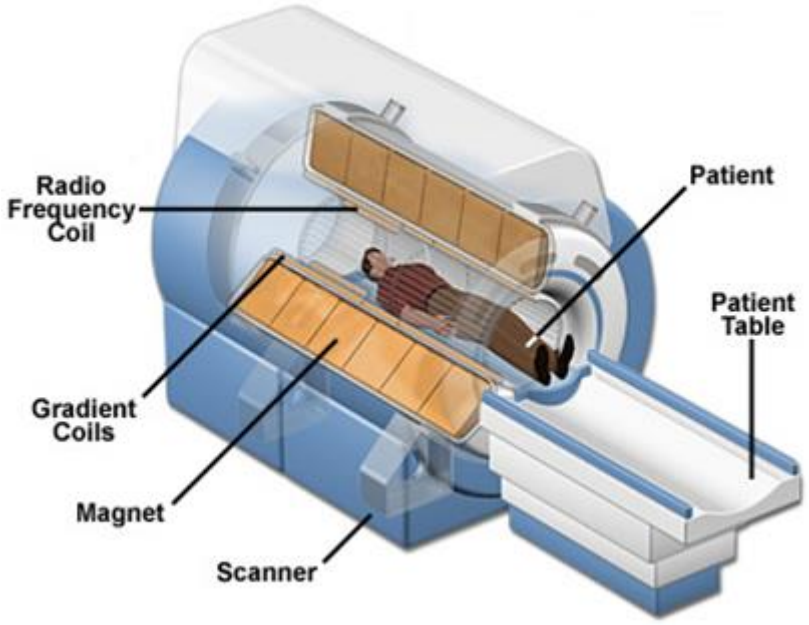


Figure 4.9 The MRI scanner. Retreived from: https://nationalmaglab.org/education/magnet-academy/learn-the-basics/stories/mri-a-guided-tour

The primary component is a superconducting electromagnet in the bore of the scanner which produces a high strength magnetic field (B_o), aligned with the

longitudinal plane of the MRI scanner (z-axis). The strength of the magnet is measured in tesla (T) and is a coil made of niobium-titanium wire which is immersed in liquid helium to ensure the temperature remains near to absolute zero, at 4.2 Kelvin. At this extremely cold temperature, the niobium-titanium wire's electrical resistance is negligible and it becomes superconducting, producing a strong and stable magnetic field.

The RF transmitter and receiver coils are aligned in the transverse plane (x-y), orthogonal to the B_o field. The RF transmitter coils generate radiofrequency pulses at the resonant frequency (Larmor frequency) of the protons in the tissue. The energy of the RF pulses is absorbed and then re-emitted by protons, and the RF receiver coils detect this signal. RF coils with overlapping fields in specialised arrays are used in structural and functional MRI to provide whole brain coverage.

Three sets of gradient coils create a secondary magnetic field which modifies the strength of B_o linearly in three orthogonal directions (x, y and z). These gradient coils enable localisation of the MRI signal in order to generate images. As protons within a tissue re-emit the resonance frequencies simultaneously, manipulating the magnetic field causes protons to absorb and re-remit energy as a function of position in the sampled tissue. By modifying the three sets of gradient coils during a scanning sequence across the z, x and y axes, slice selection, thickness and the source of a signal within a slice can be determined.

The last crucial components of the MRI scanner are shimming coils. Shimming coils produce small additional magnetic fields to compensate for inhomogeneities present in B_o. These compensatory fields ensure that B_o is as close to homogenous as

possible across the region of interest, allowing high resolution images to be obtained.

4.8.2 Principles of MRI

For brain imaging, MRI scanning relies on the detection of hydrogen protons (^{1}H) within human tissue. Isotopes of atoms with unequal numbers of nucleons (that is odd numbers of protons, or neutrons, or odd numbers of both) have nuclear spin and, as hydrogen atoms consist of single proton in the nucleus, this includes hydrogen nuclei. Nuclear spin is the self-rotation of a nucleus around its own central axis, creating a small magnetic field which is influenced by external magnetic fields, such as that of the MRI scanner. This small magnetic field creates a further magnetic field perpendicular to it. This causes the nucleus to undergo a phenomenon known as ‘magnetic moment’. Protons have a large magnetic moment, giving them relatively high sensitivity. Furthermore, protons are naturally abundant in comparison to other nuclei and are present in all tissues in the form of water, fat and other biological materials, making them the ideal nucleus for MRI studies.

Under normal conditions, the proton spins within hydrogen atoms will be randomly aligned. The high strength of the magnetic field within the MRI scanner, however, forces the nuclear spins to align with the external magnetic field (B_o), either parallel or anti-parallel. This is influenced by the strength of B_o and the temperature, with stronger magnetic fields and lower temperatures causing relatively more proton spins to be in the lower energy parallel state, creating a population difference. The net magnetization can then be manipulated by radio frequency pulses and the gradient coils, and subsequently detected by the receiver coils. The rotational motion of the nuclei as they align with the main magnetic field is called precession and the

frequency of the spinning rate (Larmor frequency) is calculated using the following formula, where ω is the Larmor frequency; γ is the gyromagnetic ratio; and B is the strength of the magnetic field:

$$\omega_o = \gamma B_o$$

In the MRI scanner, RF transmitter coils emit pulses of energy at the Larmor frequency of the precessing hydrogen nuclei which is absorbed. As this energy is absorbed, the amplitude of transverse magnetisation increases and once the RF is removed, the nuclei relax back to their normal state. The rate of the relaxation processes vary according to tissue type, enabling the identification of anatomical structures.

4.8.2.1 Signal decay

The relaxation occurs in two ways: longitudinal relaxation (T1 recovery); and transverse relaxation (T2 decay). In T1 recovery, the nuclei lose the energy obtained to their molecular surroundings and the rate at which this energy is lost depends on the composition of the tissue. Molecules are constantly moving, tumbling with certain rotational correlation times, causing fluctuations in the local magnetic field. If molecular tumbling is occurring at a rate close to the Larmor frequency (for example, in lipids), energy is released rapidly from the protons to their surroundings. If tumbling happens at a lower or higher rate, T1 recovery will be slower (for example, in solid tissue or cerebrospinal fluid, respectively). Tailoring the repetition time (TR) between RF pulses according to differential longitudinal relaxation rates ('T1 weighting'), sensitizes MRI to different tissue types.

In T2 decay, the nuclei spins dephase, resulting in a net loss of magnetisation in the x-y plane. In comparison to T1 recovery, T2 decay is rapid. Dephasing is caused by

both inhomogeneities in the internal magnetic field (interactions between neighbouring molecules and tissue) and inhomogeneities in the external magnetic field (B_o). Dephasing due to the variations in the internal magnetic field also depends on the rate of molecular tumbling. Protons in slow tumbling molecules dephase more rapidly than those in fast tumbling molecules. On the other hand, dephasing due to variations in the external magnetic field of the scanner is known as free induction decay (T2* decay). T2* decay includes influences such as shimming imperfections in the instrument setup and magnetic susceptibility changes across the tissue and it is, therefore, more rapid than natural T2 decay alone. However, as the effects are predictable and constant, they can be corrected using spin-echo imaging.

4.8.2.2 Spin-echo imaging sequences

In a spin-echo imaging sequence, refocusing RF pulses are emitted to flip the spin vectors at 180º so they swing back into alignment. This reverses the T2 decay generating an echo spin as the vectors move back into phase. Two key timing parameters can be manipulated in spin-echo imaging: the time between the initial RF pulse and signal measurement ('echo time' (TE)); and the repetition time (TR), which refers to the time between repeated refocusing RF pulses. Varying these timing parameters will affect whether the signal intensity is mainly due to T1, T2 or T2* relaxation. For example, the TE and TR are shorter in T1 weighted images to maximise the signal intensity from T1 relaxation. Oppositely, lengthening these parameters in T2 weighted images maximises the signal from T2 decay.

The final MR images provide macroscopic information about the underlying tissues in the brain. In T1 weighted imaging, which is utilised in this study, each voxel represents thousands of neurons and other supporting structures and typically has a

spatial resolution of 1mm³. White matter is primarily composed of myelinated axons, glial cells, and extracellular space; whereas grey matter consists of neuronal cell bodies, dendrites, myelinated and unmyelinated axons, glial cells, blood vessels and extracellular space (Mills & Tamnes, 2014). Although changes in underlying neurobiological processes and microstructure observed in longitudinal MR imaging cannot be directly determined using MRI; it has been proposed to reflect changes in myelination, dendritic branching/pruning, synaptogenesis, spine formation/elimination, gliogenesis and angiogenesis (Tardif et al, 2016).

4.8.2.3 Echo planar imaging

Although conventional spin-echo sequences are suitable for the acquisition of structural MRI (sMRI) data, they are not rapid enough to acquire all-phase encoded steps required for functional MRI (fMRI). To enable the acquisition of individual slices across the whole brain in fMRI, single-shot imaging, known as echo planar imaging (EPI), is used. In EPI, all data is acquired from the signal generated from one RF pulse which is followed by a train of spin-echoes which each encode a different phase. Although this technique provides low resolution images in comparison to standard MR images, due to the poorer spatial resolution, the temporal resolution is far superior, enabling whole brain data to be acquired in around two to three seconds (Buxton, 2002).

4.8.3 Principles of fMRI

The previous section detailed the use of MR in the acquisition of structural images, fMRI, on the other hand, is an experimental tool designed to map neural activity, either at rest or in response to a cognitive task. The following section will provide a brief overview of the physiological basis of this technique.

4.8.3.1 Neurovascular coupling

Fluctuating electric and electromagnetic fields can be measured by placing electrodes on the scalp, such as in electroencephalograms, however, the locational information provided by these techniques is not sufficient enough to produce detailed maps of functional activity. Due to this, fMRI attempts to measure neural activity by using a surrogate marker, cerebral blood flow (CBF). CBF continually supplies the working brain with oxygen and glucose to enable the generation of electric activity, which is necessary for neuronal signaling.

Neurons are densely interconnected and communicate via action potentials (electrical signals) which propagate through their axons, branching into axonal terminals and releasing neurotransmitters at synapses with neighbouring neurons. At rest, neurons have an electrical potential difference across their membranes, creating a more negative potential inside the cell. In the extracellular space, there is a higher concentration of sodium (Na^+) and calcium (Ca^{++}) ions, whereas intracellularly, there is a higher abundance of potassium ions (K^+) and this concentration gradient is maintained by ion channels in the neuronal membrane.

When a neuron receives synaptic input, depolarisation occurs and the voltage-gated sodium ion channels open, causing an influx of Na^+ ions. This changes the intracellular potential to a more positive charge. An action potential is produced when the threshold potential is reached and this small current travels down the axon, opening Na^+ channels in the next patch. At a synapse, the action potential triggers membrane permeability to Ca^{++} ions, causing vesicles to merge with the cell membrane and release the neurotransmitters they contain into the synaptic gap. These then bind to receptors on the postsynaptic terminal and, depending on

whether the neurotransmitter is excitatory (for example, glutamate) or inhibitory (for example, gamma-aminobutyric acid), the electrical potential of the next neuron is either increased or decreased.

This process, and the mechanisms that maintain the concentration gradient at rest and restore the electrical potentials following an action potential, require oxygen and glucose, supplied via CBF, which are metabolised to provide energy. When there is an increase in neuronal activity, the demand for oxygen and glucose rises, leading to an increase in blood flow to the local area. This is brought about by the release of the excitatory neurotransmitter glutamate which causes the release of vasodilators.

4.8.3.2 Blood-oxygen level dependent fMRI

Immediately following increased neuronal activity, there is an initial dip in blood oxygenation, followed by a rapid increase in blood flow which leads to an overcompensation to meet the oxygen demands. As blood volume rises, so does the proportion of oxygenated haemoglobin (HbO2) in comparison to deoxygenated haemoglobin (dHb), peaking at around four to six seconds before returning to baseline. This sequence of events is described as the haemodynamic response function (HRF) and the changes in oxygenation is what is used in Blood-Oxygen Level Dependent (BOLD) fMRI to create images which represent neuronal activity.

When oxygenated, haemoglobin has no unpaired nucleons, meaning it has no nuclear spin and is diamagnetic. When the oxygen is released, the deoxyhaemoglobin formed has four unpaired electrons and becomes paramagnetic, creating a magnetic moment. The presence of dHb causes local distortions in the magnetic field of B_o and these increase spin dephasing, shortening T2* decay as the amount of dHb rises. Due to this, deoxygenated blood will give a lower T2* signal

than oxygenated blood and it is this BOLD contrast that is exploited in EPI. However, the BOLD signal is a relative measure as T2* decay varies at different voxels due to differences in magnetic susceptibility for structures across the brain. To account for these differences, fMRI task-based experimental designs need to compare signal intensity in a cognitive task to that of a baseline task (for example, block or event-related designs).

4.9 Chapter 2: Longitudinal structural MRI data acquisition, pre-processing and analysis

The participants in the study who were able to be scanned, completed MRIs at three time points (T2, T3 and T4), each six weeks apart, although there was some within and across participant variability due to scheduling restrictions. Structural T1-weighted whole brain images were acquired prior to functional imaging and all scans were completed on the same 3.0 Tesla Siemens Trio MRI scanner (Siemens Medical Systems, Erlangen, Germany), with a standard 20-channel head coil. A T1-weighted 3D modified driven equilibrium Fourier transform sequence was used, which produced 176 contiguous sagittal slices with a 256 x 224 matrix (resolution = $1mm^3$; repetition time/echo time/inversion time = 7.92/2.48/910ms; flip angle = 16°). All T1-weighted scans were analysed using the software programme Statistical Parametric Mapping (SPM12) in MATLAB 2018b.

4.9.1 Lesion identification

Historically, the gold standard for lesion identification was for an expert to manually delineate abnormal brain tissue, however, this technique is time consuming and can be affected by subjective bias (Wilke, de Haan, Juenger & Karnath, 2011; Ashburner & Friston, 2000). To overcome these issues, automated computer algorithms can be

used that give higher inter-rater and intra-rater reliability. To identify patient lesions in the present study, and for preprocessing prior to statistical analysis, the Automated Lesion Identification (ALI) toolbox in SPM12 was used (Seghier, Ramlackhansingh, Crinion, Leff, & Price, 2008).

ALI is a modified segmentation procedure which segments healthy and damaged brain tissue by comparing patient volumes to those of healthy controls. It is based on the assumption that the damaged tissue (the lesion) will be comprised of atypical, outlier, voxels which do not correspond with the conventional tissue types (grey matter (GM), white matter (WM), or cerebrospinal fluid), and have extreme intensities that are not within a normal range. To avoid misclassification in the segmentation process, these atypical voxels are classified as an extra tissue class, enabling identification of the lesion.

The raw T1-weighted baseline scans (collected at time point T2) were entered into the ALI toolbox using default parameters to produce a normalised binary lesion image for each patient (Seghier et al, 2008). These binary images were manually compared to the patients' raw scans and any regions which were incorrectly identified as damaged tissue were removed using MRIcron (Rorden & Brett, 2000). The resulting binary images were combined to generate a group lesion overlap map, displayed in Chapter 2.

4.9.2 Voxel-Based Morphometry

Voxel-based morphometry (VBM) is an automated, mass-univariate, technique which is used to investigate volumetric differences across the whole brain, on a voxel-by-voxel basis (Ashburner & Friston, 2000; Mechelli, Price, Friston, & Ashburner, 2005). It is considered an un-biased and objective technique as the whole brain is

examined, rather than pre-selected structures. The aim of VBM is to identify small-scale differences in the local grey and white matter, whilst omitting differences in large-scale gross anatomy, either between individuals or across individuals over time. Although the exact nature of these differences continues to be speculative, it has been suggested that experience-dependent changes may reflect changes in neuronal size, axonal or dendritic arborisation, or neuropil (Mechelli, Price, Friston, & Ashburner, 2005). For GM specifically, decreases in volume could indicate neuronal cell loss or atrophy; and increases may signify thickened GM (Seghier, Ramsden, Lim, Leff, & Price, 2014; Hervais-Adelman, Moser-Mercer, Murray, & Golestani, 2017).

4.9.2.1 Longitudinal VBM

In the present study, serial longitudinal registration was employed to investigate changes in GM and WM from pre-therapy (T2-T3) to therapy (T3-T4). This type of analysis has been found to be sensitive in identifying subtle changes over time and has been used with patients with primary progressive aphasia to examine atrophy in language regions (Mandelli et al, 2016; Santos-Santos et al, 2016). The exact procedure used in this study was also used to investigate changes in the brain following the completion of a computerised comprehension therapy in people with post-stroke aphasia (Fleming et al, 2020).

Although the within-subject design of longitudinal serial imaging increases its power, it is also affected by additive bias in which changes in volume are biased towards the first time interval. To reduce this bias, SPM combines diffeomorphic registration, intensity non-uniformity correction and rigidbody registration, which were used in the iTalkBetter study to create divergence maps for each participant for pre-therapy (T2-

T3) and therapy blocks (T3-T4) (Ashburner, 2013). Within these maps, values represent either compression ('loss') (<1), or expansion ('gain') (>0), at each voxel. The divergence maps were combined with participants' average GM and WM images to provide a single difference image for both GM and WM (Figure 4.10).

To produce these difference images, the following steps were completed:

1.) The three raw T1 weighted images (T2, T3 and T4) were loaded into the serial longitudinal registration toolbox, using default parameters, to produce an average image and three divergence maps. The time of the scan was entered as a decimal of a year.
2.) The average image was segmented using the ALI toolbox, again using default parameters, to produce average images for GM and WM.
3.) The divergence maps for GM and WM, separately, were subtracted from one another, generating two divergence maps for each which corresponded to change over pre-therapy (T3-T2) and change over therapy (T4-T3). These were divided by the time interval between scans for each participant, entered as a decimal of a year, to account for variabilities in scanning schedules.
4.) The resulting divergence maps were multiplied by the participants' average GM and WM images, from step 2, to produce probabilistic 'change' maps.
5.) The pre-therapy probabilistic change maps were subtracted from the therapy change maps to produce two images for each participant: one for GM and one for WM. These represented change over therapy, more than change over pre-therapy.
6.) Images were then spatially normalised into the same stereotactic space to allow for group level analysis, which in this case was Montreal Neurological

Institute (MNI) space, using the deformation field from the normalised average scan image.

7.) The normalised images were then smoothed with an isotropic Gaussian kernel of 8mm at full-width half maximum to account for any anatomical variability not resolved in spatial normalisation and to improve the signal-to-noise ratio.

A general linear model was used following the pre-processing steps above. This produced 3D statistical maps, representing a statistical test at each voxel, which related to a change in the volume of GM or WM. Correction for multiple comparisons, using Gaussian random field theory, was applied to reduce the risk of false positive results (Type 1 error). The correction level required is dependent on the smoothness of the images and the number of resels (resolution elements) is calculated using the Euler statistic from the 3D statistical map. In the present study, correction for multiple comparisons was completed by reporting significant clusters at the Family Wise Error (FWE) $p<0.05$ threshold (statistical voxel-level threshold was set at $p<0.001$ uncorrected).

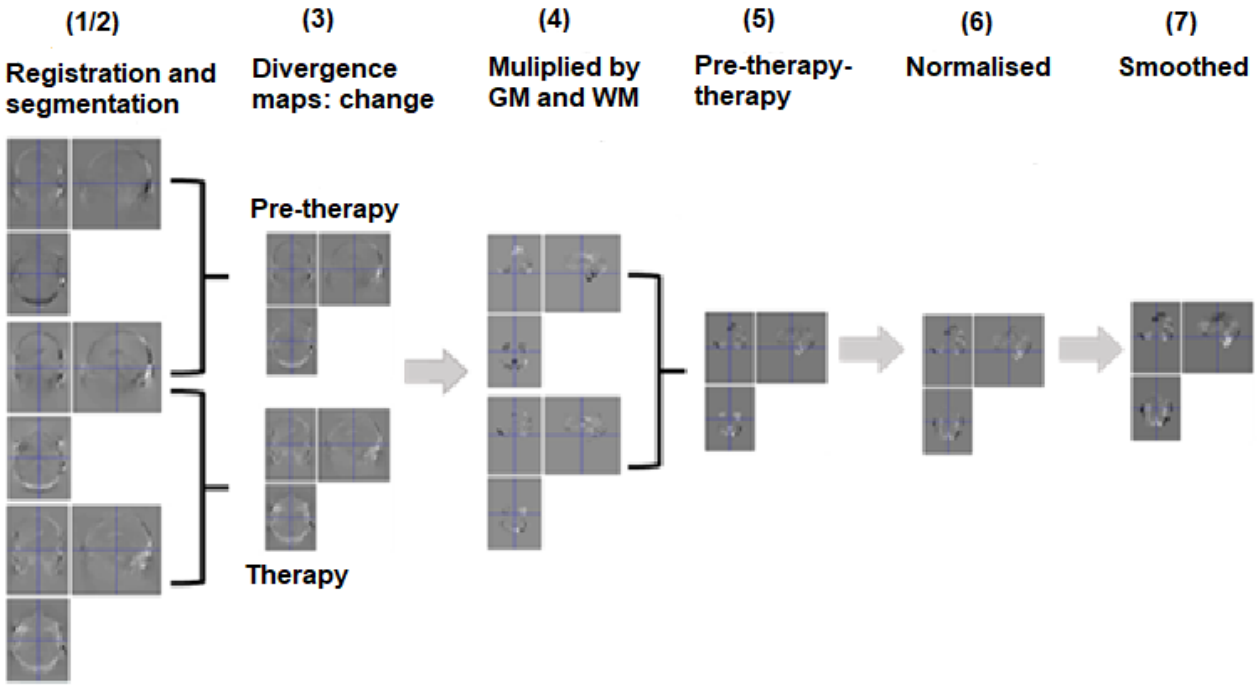


Figure 4.10 Longitudinal VBM pre-processing pipeline for one participant.

4.10 Chapter 3: Longitudinal functional MRI data acquisition, pre-processing and analysis

Task-based functional images were attained immediately following the T1 weighted structural images, at each of the three time points (T2, T3 and T4). A gradient EPI sequence was used (matrix size=64×64; repetition time/echo time = 3080/30ms; flip angle = 90°; field of view = 192×192; slice thickness = 2mm, inter-slice gap = 1mm) and scans were acquired on the same scanner as the structural scan with the same standard 20-channel head coil. Each task-based functional run consisted of 66 volumes per time series. This included five 'dummy scans' to allow for magnetisation to reach equilibrium. The fMRI task-based data was pre-processed using SPM12 in MATLAB R2015a.

Resting state fMRI imaging was also acquired after the task-based fMRI, using the same EPI acquisition sequence as above. 100 whole-brain volumes were attained

(comprising of 44 slices each) and, as with the task-based fMRI, the first five scans were discarded as ‘dummy scans’. The resting state data were not analysed as part of my thesis.

4.10.1 Tasked-based experimental design

The task-based fMRI experiment followed an existing protocol as the data was collected as part of the Predicting Language Outcomes and Recovery After Stroke (PLORAS) research project (Sanjuán et al, 2014). The experiment was a block design and consisted of five tasks: two finger press response tasks (semantic picture decision and semantic auditory decision) and three overt speech tasks (object naming, verb production and sentence production). Each task had 20 pictures or auditory stimuli which were made up of two items, either people, animals or objects, with names of one to four syllables (see Figure 4.11).

Figure 4.11 Task-based fMRI picture examples. A: Semantic picture decision; B: Object naming; C: Verb production; D: Sentence production.

In the semantic decision tasks, participants were required to decide whether two pictures (picture decision) or two spoken words (auditory decision) were semantically related (for example, ‘pirate and boat’) or semantically unrelated (for example, ‘frog and plate’) via a button press.

In the object naming task, participants had to name aloud two unrelated items in a picture (for example, ‘goat and train’). For both verb production and sentence production, the pictures presented depicted two items interacting with one another. In the verb task, participants were required to produce only the verb to describe what was happening (for example, ‘jumping’). For the sentence task, participants had to describe what was happening using a full sentence (for example, ‘the nurse is drinking from the glass’). To constrain the responses for these two tasks, only four verbs were incorporated into the pictures (eating, drinking, jumping and falling).

All five tasks and the stimuli within the tasks were presented in the same order for each participant and at each time point (Table 4.4).

Task order	Stimuli	Task instructions	Response
1. Semantic picture decision	2 pictures	Are the objects/ animals related or unrelated	Finger press
2. Object naming	2 pictures	Name the two objects/ animals	Speech
3. Verb production	Event picture	Name the verb	Speech
4. Sentence production	Event picture	Describe what is happening	Speech
5. Semantic auditory decision	2 auditory words	Are the words are related or unrelated	Finger press

Table 4.4 The five fMRI tasks.

4.10.1.1 Procedure

Prior to each scanning session, participants were trained on the five tasks using stimuli not displayed in the scanner. In the scanner, each task was presented in a

separate scan run and the 20 items in a task were presented in four blocks of five, at a rate of one every 5 seconds. Each block lasted for 20 seconds to allow for sufficient time for the BOLD response to peak. At the end of a block was a 16 second rest block to allow activation, and the proportion of oxygenated and deoxygenated blood, to return to baseline. During these breaks, a fixation cross was displayed on the screen. Before the next block started, written instructions (for example, 'Name the verb') were given.

Picture stimuli were displayed for 2.5 seconds (followed by a 3.5 second fixation cross) on a projector that participants viewed via an adjustable mirror on the head coil. Auditory stimuli were presented through MRI compatible headphones (MR Confon, Magdeburg, Germany) for 1.75-2.5 seconds, during which, a fixation cross was shown. Speech responses were recorded via a noise-cancelling MRI microphone (FOMRI IIITM Optoacoustics) and were also transcribed manually online. In the semantic decision tasks, an MRI compatible button box was used to record responses (participants used two fingers on the same hand to press one of two buttons).

Each scanning session lasted approximately 1 hour 30 minutes. This included pre-scanning training, positioning the participant in the scanner and completing structural and functional imaging.

4.10.1.2 Behavioural scoring

The audio recordings were used to verify the online transcribed speech responses and the button box responses were received from the recorded data. Trials were marked as either correct (matched the target), incorrect (did not match the target), or no response (no output or verbal fillers such as 'erm' and 'I don't know'). Additionally,

for the object naming task, responses were marked as correct if a participant said at least one of the target words. For the sentence production task, responses were marked as correct as long as at least two of the arguments in the sentence (player, verb or object) were produced correctly.

4.10.2 fMRI pre-processing

In the scanner, measures were taken to minimise head movement, for example, using padding around the participant's head, but a small amount of movement was unavoidable, particularly for speech production tasks. As a result, motion correction was necessary to ensure the location of a single voxel within the field of view was constant, and any changes in signal intensity were due to a change in the BOLD response. To correct for movement in the pre-processing steps, functional images were spatially realigned to the first EPI volume, to bring all of the data into the same orientation, and un-warped to account for residual movement related variance induced by susceptibility-by-movement interactions, using the unwarp toolbox in SPM12.

Motion parameters can also be added as covariates of no interest in the first level design matrix so that signal changes that correlate with movement can be ignored. However, if head movement is associated with the condition of interest (such as in the speech production tasks used in this study), this can result in a loss of real changes in signal. Due to this, no motion parameters were added to the design matrices in this analysis.

Following motion correction, the structural T1 image was co-registered to the mean EPI image and spatially normalised into standard MNI space using the unified normalisation-segmentation tool in SPM12. Using the deformation parameters

acquired during the normalisation of the T1 structural image, EPI scans were then spatially normalised to MNI space and smoothed using an isotropic Gaussian Kernel at 8mm full-width-half-maximum.

4.10.3 First Level Analysis

In the first level design matrix, all five of the task-based functional pre-processed images from the three time points (T2, T3 and T4) for each individual were entered into a fixed-effect general linear model (GLM) (Friston et al, 1994). For each of the five tasks (two non-verbal button box tasks and three speech tasks), the stimulus onset times were modelled as single, spike events, alongside response durations and reaction times, with four distinct regressors for each task: 1) instructions; 2) correct responses; 3) incorrect responses; and 4) no responses. However, due to the severity of aphasia for some participants in the study, for the three speech tasks distinctions were not made between ‘correct’ and ‘incorrect’ responses. Instead, these two response types were collapsed together and coded as ‘speech responses’ to enable all participants to be included in the analysis. Stimulus functions were then convolved with a canonical haemodynamic response function. To remove low frequency confounds (noise) caused by ‘scanner drift’ (linear changes over time in signal intensity), temporal filtering was performed using a high-pass filter of a set of discrete cosine basis functions with a cut-off of 128s.

Due to limitations in analysing the task-based fMRI data (as scans were collected, pre-processed and analysed as part of the PLORAS project), as opposed to the structural imaging analysis which compared change over pre-therapy to change over therapy, within-subject contrasts were created which represented change in functional activation at post-therapy (T4) in comparison to an average of functional

activation at pre-therapy (T2 and T3). For this, the speech responses for the three speech production tasks were contrasted against an implicit baseline (rest), and therapy effects were modelled by contrasting the pre-therapy time points with the post-therapy time point. For each participant, all data was entered into one first level GLM and the contrast for each task was modelled as: (T2) 0 -1 0 0; (T3) 0 -1 0 0; (T4) 0 2 0 0 (Figure 4.12).

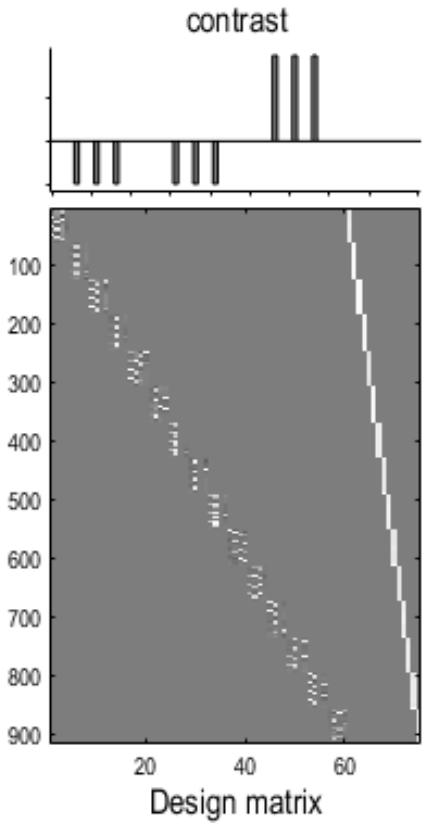


Figure 4.12 First level design matrix for each participant.

4.10.4 Second level analysis

To investigate functional changes from pre-therapy to post-therapy at the group level, the contrast images produced from the first level analysis for all participants were entered into a one sample t-test in SPM12. Both increases and decreases in activation at post-therapy in comparison to pre-therapy were separately analysed,

using the contrast 0 1 for increases; and 0 -1 for decreases. As with the structural analysis, the statistical threshold was set at $p<0.001$ uncorrected, and significant clusters are reported at $p<0.05$ using family-wise error (FWE) correction to control for multiple comparisons across the whole brain.

5 Results

5.1 Chapter 1: Investigating the behavioural response to iTalkBetter therapy in persons with chronic aphasia.

Aim 1: To investigate whether iTalkBetter can improve single word retrieval in persons with chronic aphasia.

Aim 2: To investigate if improvements in word retrieval in a single word naming test generalise to improvements in word retrieval in a spoken picture description task.

Aim 3: To investigate if baseline factors can explain response to therapy.

Hypotheses:

(1) At the group level, participants will significantly improve their retrieval of trained lexical items following iTalkBetter therapy and this improvement will not generalise to (psycholinguistically matched) untrained items.
(2) At the group level, participants will significantly improve their ability to retrieve both trained concrete and trained abstract lexical items.
(3) Following iTalkBetter therapy, participants' will significantly improve their retrieval of trained lexical items in the spoken picture description task.
(4) A combination of demographic and baseline behavioural factors will explain individual response to the iTalkBetter therapy.

5.1.1 Experimental procedures

5.1.1.1 Participants

Data from 27 participants are included in the following analyses. Key demographic and baseline behavioural data are presented in the Methods.

5.1.1.2 Design

A repeated measures design was used, with five testing time points (T1-T5). A detailed description of the study design can be found in the Methods.

5.1.1.3 Data pre-processing

iTalkBetter therapy data

Dose is the total time participants spent completing iTalkBetter therapy over the six-week therapy block (rounded to hours). Exposure is the number of exposures to trained lexical items, which corresponds to the number of cycles of therapy participants completed. As the first seven participants were given a larger therapy corpus, and therefore, had less exposure to trained items, dose and exposure are analysed separately in the following sections. Both dose and exposure data were recorded automatically by the app. This data was entered manually into Excel and statistical analyses were completed in SPSS25.

Outcome measures and baseline assessments

All outcome measures (across each of the five time points) and baseline assessments were scored by hand. As above, this data was entered into Excel and analysed using SPSS25.

5.1.1.4 Statistical analyses

Hypotheses 1 and 2: To investigate whether iTalkBetter improved the retrieval of single words, repeated measures ANOVAs compared change on the main outcome measure (the Word Retrieval Test) over the six-week therapy block with change over the six week pre-therapy block.

Hypothesis 3: A repeated measures ANOVA was completed to ascertain if there were improvements in word retrieval in the SPD task.

Hypothesis 4: Simple correlational analyses were employed to examine whether therapy dose (hours), exposure (practice of a single trained lexical item) or baseline severity could predict response to therapy. To investigate if a combination of baseline factors could explain the variability in response to therapy, automatic linear modelling (ALM) was used.

5.1.2 Results

Definitions of terms used in this section:

Trained lexical items: 110 lexical items trained in the iTalkBetter therapy and tested in the WRT at each time point.

Untrained lexical items: 110 matched lexical items not trained in the iTalkBetter therapy but tested in the WRT at each time point.

Core lexical items: Lexical items trained in the therapy but not tested in the WRT (780 for the first seven participants, and 110 for the following 20 participants).

Therapy corpus: All lexical items trained in therapy, including both trained lexical items and core lexical items (890 for the first seven participants, and 220 for the following 20 participants).

Dose: The number of hours of therapy completed.

Exposure: Exposures to trained lexical items within the iTalkBetter therapy, not including practice of the same word at multiple cue depths.

Lexical item trials: Each individual practice in therapy of trained lexical items at all cue depths (see section 4.5.1 in Methods).

5.1.2.1 Therapy dose and exposure

Figure 5.1 displays the number of hours of therapy achieved and the number of exposures to trained lexical items over the iTalkBetter therapy block for each participant. On average, participants completed 45 hours of therapy over six weeks (SD=16), and were exposed to each lexical item an average of 27 times (SD=16). As expected, there was a significant positive correlation between dose and exposure, r=.39, p=0.04. As as illustrated in Figure 5.1, however, the first seven participants received significantly less trained lexical item exposures (M=9, SD=5), than the following 20 participants (M=32, SD=14), U=3, p=<0.001, despite there being no significant difference in dose between the two groups, U=59, p=0.57. This was because the first seven participants completed a much larger therapy corpus (890 lexical items as opposed to 220 lexical items) and, therefore, completed less therapy cycles over the six weeks of therapy. During the iTalkBetter therapy, participants completed, on average, 17,538 individual lexical item trials (SD=6,262), and this was highly correlated with dose, r=.99, p=<0.001.

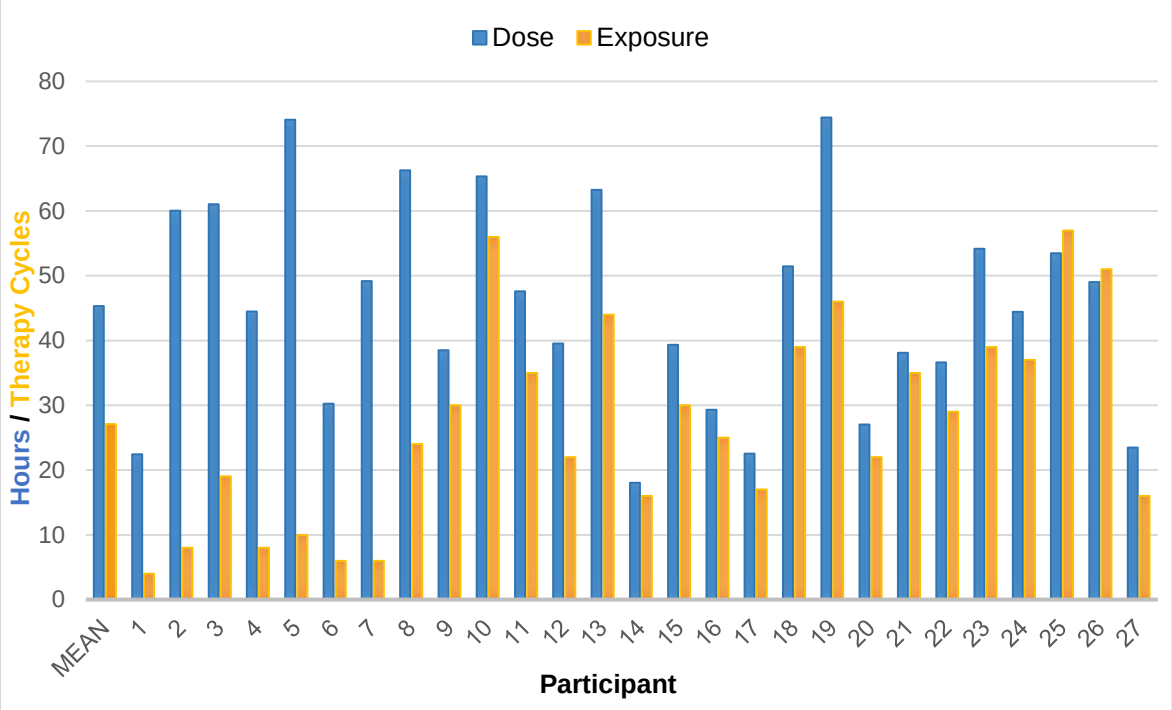


Figure 5.1 Dose and exposure over the six-week therapy block. For each participant (N=27), blue bars are the number of hours of therapy completed (dose) and yellow bars are the number of exposures to trained lexical items. The first two bars represent the group mean for dose and exposure.

5.1.2.2 Baseline performance

Word Retrieval Test

Participants' performance at baseline (T1) on the WRT was 32% (SD=15.77), significantly below that of 38 aged-matched healthy controls (M=87%, SD=9.15), $t(38.5)=16.07$, p=<0.001 (equal variances not assumed). At the group level, there was a small but significant improvement on the WRT from T1 to T2 (2.95%), $t(26)=2.91$, p=0.007, as well as between T2 and T3 (3.16%), $t(26)=2.86$, p=0.009. There was, however, no significant difference in change between T1 to T2; and T2 to T3, $t(26)=0.13$, p=0.9, indicating this was a stable improvement over the three pre-therapy time points.

Linguistic and cognitive measures

Baseline performance on naming and repetition is presented in the Methods and scores for further key linguistic and cognitive measures are shown in Table 5.1, alongside cut-off criteria for each test where available

All participants scored below the aphasia cut-off on the 'Object Naming' subtest of the CAT, demonstrating their difficulties with word finding. Additionally, all participants were impaired on comprehension of spoken and written sentences, and all but one (P25) on single word reading, indicating that language deficits were present across both input and output modalities. All participants, however, had poorer single word naming in comparison to spoken single word comprehension, and, although variable, no participant had a severe impairment in single, real-word repetition.

All participants bar one (P24) also scored below the cut-off on the non-word repetition subtest (subtest 9) of the PALPA, suggesting impairments with phonological input-to-output conversion, and all but one participant (P14) scored below the aphasia cut-off on the spoken picture description of the CAT. Both P24 and P14 were the only participants who were not impaired on Digit Span.

Of the 20 participants who were able to complete the cognitive assessments that required in-person testing, all scored below the norm on the 4-Way Weigl (4-WW) and half scored below the cut-off criteria on the Cattell Culture Fair Intelligence Test (CFIT) and the Sustained Attention to Response Task (SART) (omissions of responses on 'Go trials'). This indicates that these participants had non-verbal cognitive impairments in addition to their language deficits, and may have difficulties with executive functioning abilities such as, set shifting, problem solving and sustaining attention.

Table 5.1 Performance on key baseline linguistic and cognitive measures

			Spoken Comp		Written Comp								
ID	WRT (%)	SPD	Words (/30)	Sentences (/32)	Words (/30)	Sentences (/32)	Reading (/48)	Digit Span (/14)	PALPA (/80)	DVAMS (/100)	CFIT (%)	4-WW (/12)	SART-OM (/192)
P1	**35**	**9**	28	**20**	**26**	**14**	**15**	**6**	**11**	74.00	58	**8**	0
P2	**34**	**18**	27	**23**	**21**	**12**	**24**	**8**	**17**	79.70	50	**8**	2
P3	**14**	**5**	**19**	**9**	**19**	**8**	**16**	**4**	**34**	81.30	54	**1**	2
P4	**33**	**12**	**25**	**11**	30	**13**	**32**	**4**	**3**	85.30	54	**9**	4
P5	**16**	**0**	27	**16**	**21**	**14**	**5**	**4**	**30**	53.00	50	**7**	**8**
P6	**45**	**12**	26	**25**	**24**	**9**	**22**	**6**	**43**	84.40	58	**6**	0
P7	**60**	**7**	26	**22**	29	**18**	**32**	**4**	**36**	49.30	**31**	**9**	**18**
P8	**33**	**11**	**25**	**20**	**20**	**12**	**34**	**6**	**14**	81.60	50	**6**	2
P9	**26**	**3**	**23**	**8**	**20**	**11**	**28**	**0**	**4**	62.90	-	-	-
P10	**40**	**13**	**24**	**10**	**24**	**12**	**15**	**4**	**37**	61.40	-	-	-
P11	**39**	20.5	**25**	**15**	**26**	**12**	**42**	**6**	**51**	54.30	-	-	-
P12	**0**	**0**	**18**	**9**	**9**	**1**	**0**	**0**	**33**	71.43	-	-	-
P13	**19**	**4**	**12**	**16**	**5**	**4**	**40**	**8**	**54**	81.40	-	-	-
P14	**49**	**37**	**20**	**16**	**23**	**17**	**33**	10	**61**	78.60	-	-	-
P15	**28**	**4**	**20**	**22**	30	**19**	**33**	**4**	**14**	54.29	-	-	-
P16	**34**	**3**	**22**	**17**	**21**	**12**	**31**	**8**	**17**	81.43	58	**7**	**28**
P17	**26**	**3**	**19**	**10**	**19**	**11**	**31**	**6**	**56**	71.43	**30**	**5**	**13**
P18	**53**	**30**	**19**	**15**	**15**	**14**	**27**	**4**	**28**	71.43	**23**	**5**	**8**
P19	**8**	**3**	**19**	**12**	**23**	**7**	**0**	**4**	**10**	97.14	**42**	**6**	**30**
P20	**39**	**10**	**19**	**20**	**12**	**9**	**6**	**6**	**11**	100.00	**46**	**2**	2
P21	**28**	**4**	**20**	**11**	**23**	**12**	**27**	**6**	**38**	75.71	**46**	**9**	**16**
P22	**35**	**7**	28	**21**	**25**	**14**	**36**	**6**	**33**	72.86	**46**	**9**	**30**
P23	**17**	**0**	**14**	**11**	**18**	**11**	**32**	**6**	**2**	80.00	69	**7**	**25**
P24	**36**	**25**	28	**17**	**22**	**3**	**36**	10	67	22.86	**38**	**8**	**59**
P25	**62**	**5**	26	**11**	**21**	**8**	47	**6**	**52**	64.29	**15**	**4**	1

P26	**49**	**6**	27	**16**	**25**	**16**	**42**	**4**	37	61.43	81	**9**	1
P27	**4**	**2**	**22**	**20**	**19**	**12**	**0**	**4**	20	52.85	54	**6**	**48**
M	**32**	**9**	**23**	**16**	**21**	**11**	**25**	**5**	**30**	70.53	**48**	**7**	**16.17**
CO	83	33	25	27	27	23	45	8	62.5	-	<50	<10	>6

Table 5.1 Performance on key linguistic and cognitive measures. M=Mean; CO=Cut-Off criteria. Scores in bold are those below the cut-off criteria. SPD (spoken picture description), spoken comp and written comp (comprehension), reading and digit span are subtests from the CAT.

Normative values:

- WRT: Fifth percentile cut-off based on data from 38 aged-matched controls.
- CAT: aphasia cut-off scores (from the standardised assessment).
- CFIT: 1 SD below the mean based on data from 27 aged-matched controls (M=66; SD=14).
- PALPA: impairment cut-off scores (from the standardised assessment).
- 4-WW (raw score): 1 SD below the mean based on data from 23 aged-matched controls (M=11.3; SD=1.2).
- SART-OM (omissions errors): 1 SD below the mean based on data from 23 aged-matched controls (M=3.2; SD=3.6).

5.1.2.3 Therapy effects: Word Retrieval Test

Performance on the WRT at each of the five testing time-points (T1-T5) is displayed in Figure 5.2. To investigate the efficacy of iTalkBetter in improving the retrieval of trained single words on the WRT, a repeated measures ANOVA was used which compared change in accuracy over the pre-therapy block and the therapy block. There were two factors, each with two levels: time (pre-therapy (T2 to T3) and therapy (T3 to T4)); and item (trained and untrained lexical items). This analysis identified a significant interaction between time and item, F(1, 26)=10.41, p=0.003.

Follow-up paired samples t-tests revealed an item-specific learning effect as a significant change in trained lexical items from pre-therapy to post-therapy was found, *t*(26)=4.04, p=<0.001, but no significant change in untrained lexical items, *t*(26)=0.93, p=0.36. Unstandardised and standardised effect sizes for changes in word retrieval are shown in Table 5.2 for both trained and untrained lexical items, at all time-points. At the group level, participants improved their production of trained lexical items over the therapy block by 13% (SD=11.71) and this effect size was moderate to large as indicated by Cohen's *d* (0.52) and repeated measures Cohen's *d* (1.66) (Cohen's ^{d}RM). There are several ways of computing Cohen's ^{d}RM, which all acknowledge that some of the variance across the two data sets is within-subject. For this analysis, the following method of calculation was used:

(https://www.psychometrica.de/effect_size.html)

$$d = \frac{|m_1 - m_2|}{\sqrt{s_1^2 + s_2^2 - (2rs_1s_2)}}$$

Maintenance

Maintenance of the iTalkBetter therapy effects were examined using paired sample t-tests, comparing performance accuracy (%) at T4 (immediately following therapy) and T5 (12 weeks after therapy cessation), for both trained and untrained lexical items. Interestingly, performance accuracy for trained lexical items declined slightly more than for untrained lexical items (-3% and -0.2%, respectively). However, declines for both item types were small and non-significant (trained lexical items: $t(26)=1$, $p=0.33$; untrained lexical items: $t(26)=0.13$, $p=0.9$). Furthermore, performance on trained lexical items at T5 (M=47.9, SD=25.54) was significantly higher than performance on untrained lexical items at T5 (M=42.67, SD=23.33) $t(26)=4.47$, $p=<0.001$, and significantly higher than performance on trained lexical items at pre-therapy (T3) (M=37.93, SD=20.27), $t(26)=3.55$, $p=0.002$.

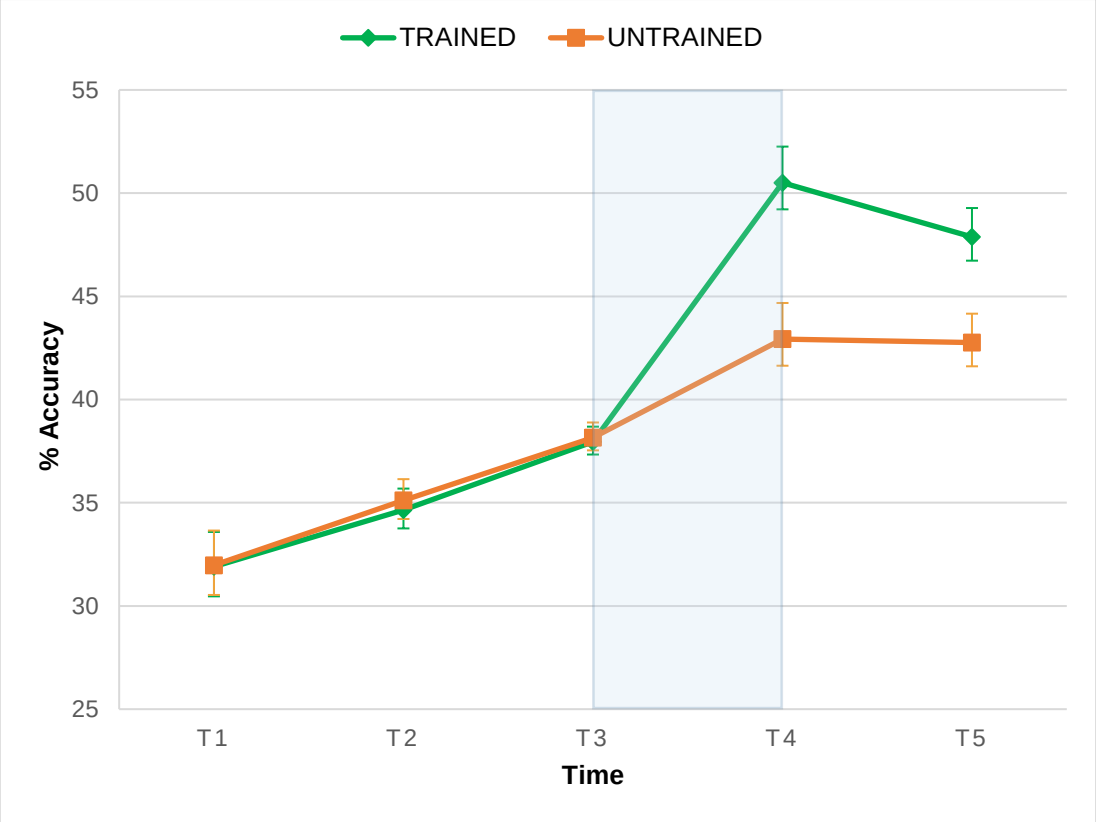


Figure 5.2 Performance (%) on the Word Retrieval Test at the five testing time-points. The green line is trained lexical items, the orange line is untrained lexical items. The blue shaded area denotes the six week therapy block. Error bars are within-subject standard error of the mean.

	Trained lexical items				Untrained lexical items			
	T1-T2	**T2-T3**	**T3-T4**	**T4-T5**	**T1-T2**	**T2-T3**	**T3-T4**	**T4-T5**
% change	2.76	3.28	12.56	-2.63	3.19	3.04	4.78	-0.17
95% CI	4.96	5.5	17.17	-0.38	5.39	5.44	7.85	2.2
Cohen's *d*	0.17	0.18	0.52	-0.1	0.19	0.16	0.21	-0.01
Cohen's ^{d}RM	0.37	0.66	1.66	-0.2	0.59	0.31	0.85	-0.01

Table 5.2 Unstandardised and standardised effect sizes for performance (%) on the Word Retrieval Test at the five testing time-points. CI=confidence interval.

Concrete and abstract lexical items

Due to the inclusion of both concrete and abstract lexical items in iTalkBetter, accuracy on the WRT for these word types were also explored using a repeated measures ANOVA, using the same design as above (performance accuracy is displayed in Figure 5.3). For this analysis, there was an additional factor of 'concreteness', which had two levels: abstract and concrete lexical items. Prior to completing this ANOVA, all of the items from the WRT were assigned to either concrete or abstract word lists using a median split based on concreteness ratings (Brysbaert, Warriner & Kuperman, 2013). A significant interaction between time and item from the main group analysis was found, but no significant interactions between time and concreteness, $F(1, 26)=0.54$, $p=0.47$, or between time, item and concreteness, $F(1, 26)=0.21$, $p=0.65$. To follow-up, post-hoc paired sample t-tests and Wilcoxon signed ranks tests (for variables which were not normally distributed: Kolmogorov-Smironov statistic $p=<0.05$) were computed.

Over the six weeks of therapy, accuracy for trained concrete and abstract items improved by 10% (SD=8.89) and 15% (SD=17.2), respectively, and these improvements were significant in comparison to the pre-therapy block (trained concrete: $t(26)=3.12$, $p=0.004$; trained abstract: $Z=2.27$, $p=0.023$). There were no significant changes in accuracy for either untrained concrete items, $Z=0.65$, $p=0.52$, or untrained abstract items, $Z=1.36$, $p=0.18$. There was also no significant difference between the improvement of trained concrete and trained abstract lexical items over the therapy block, $Z=1.7$, $p=0.09$.

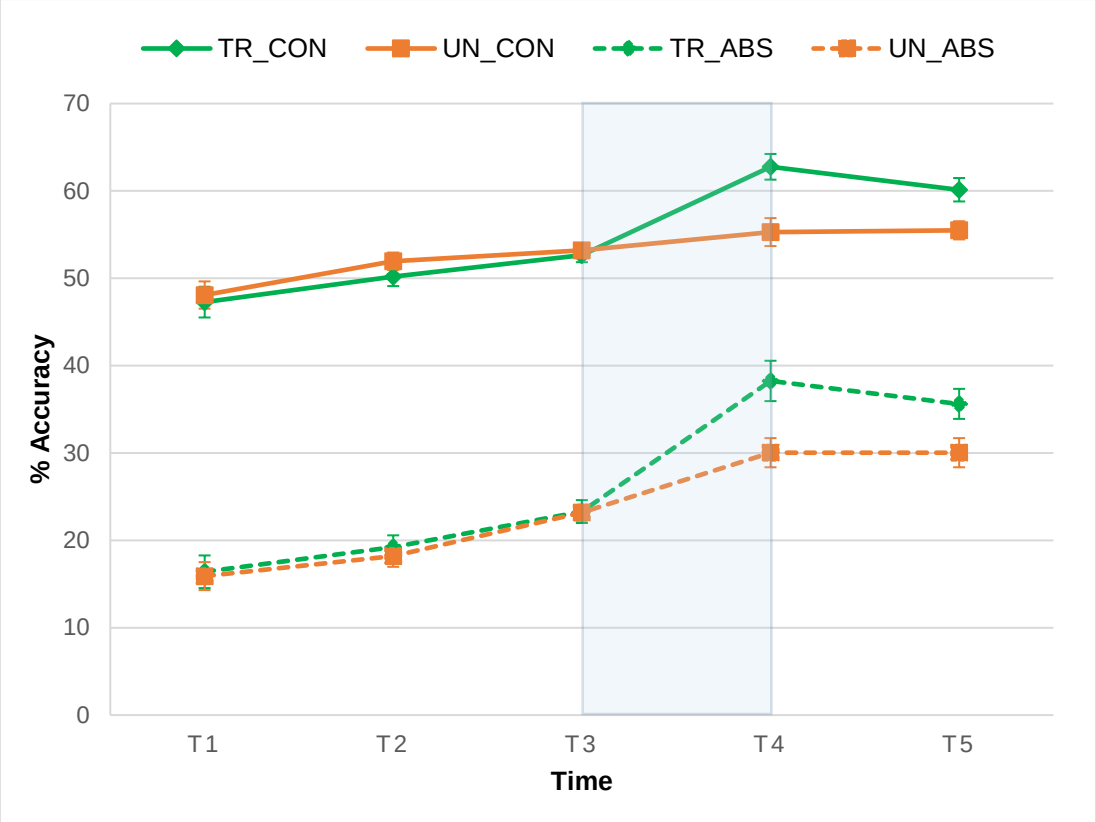


Figure 5.3 Performance (%) on the Word Retrieval Test for concrete and abstract lexical items. TR_CON=trained concrete items; UN_CON=untrained concrete items; TR_ABS=trained abstract items; UN_ABS=untrained abstract items. The blue shaded area denotes the six-week therapy block. Error bars are within-subject standard error of the mean.

Error analysis

All participants' responses in the WRT were error coded (see section 4.6.1.5 in Methods) and changes in error patterns for trained items (% of error type) from pre-therapy (T3) to post-therapy (T4) were analysed using paired samples t-tests and Wilcoxon signed ranks tests. For this analysis, certain error subtypes were collapsed together to form the following categories: no responses (including fillers and partial phonemic responses); semantic errors (including circumlocutions and synonyms); phonological errors; unrelated errors (including neologisms and perseveration); visual errors; and mixed errors.

The most common error type at pre-therapy was no response, accounting for, on average, 21% of responses (SD=21.35), followed by unrelated errors (M=17%, SD=21.59), semantic errors (M=16%, SD=8.22), visual errors (M=3%, SD=2.32), phonemic errors (M=3%, SD=3.72) and mixed errors (M=2%, SD=1.58) (Figure 5.4). At T4, production of all error types decreased at the group level, however, decreases in the following error categories were significant: no response (post-therapy: M=15.8, SD=4.29, Z=-2.95, p=0.003), semantic error (post-therapy: M=13.7, SD=1.66, *t*(26)=-2.31, p=0.03), visual error (post-therapy: M=2.22, SD=4.07, Z=-2.3, p=0.02), and phonemic error (post-therapy: M=1.55, SD=0.55, Z=-3.37, p=0.001) (Figure 5.5).

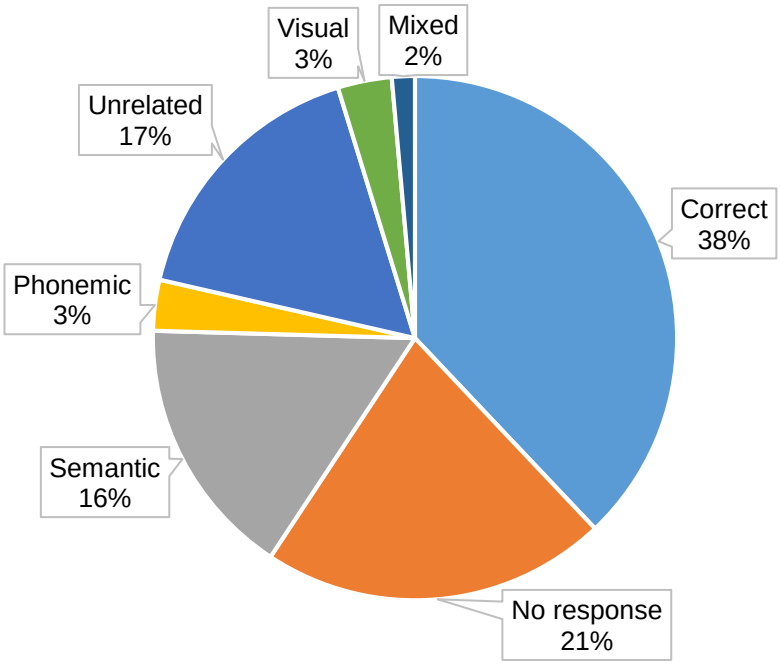


Figure 5.4 Percentage of response types at pre-therapy (T3).

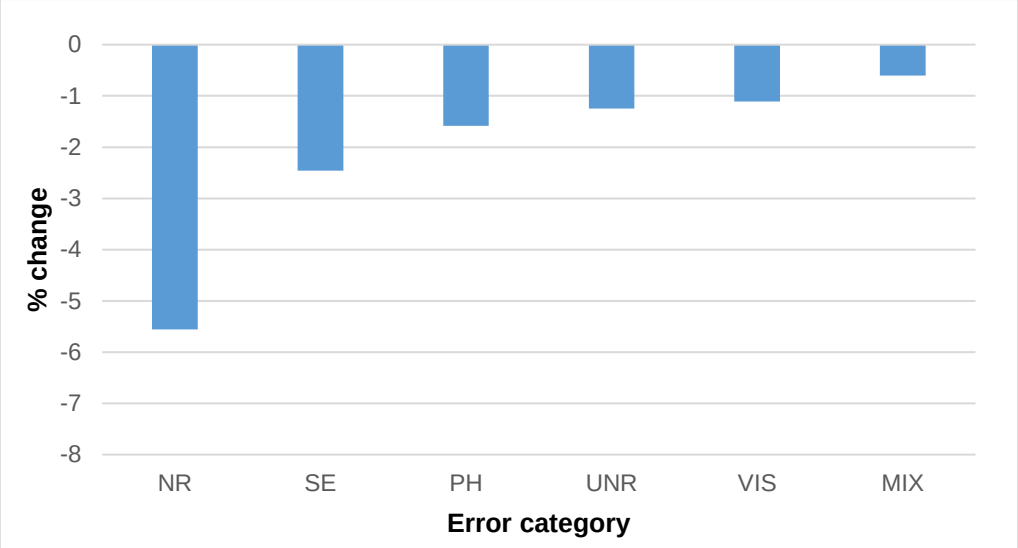


Figure 5.5 Percentage change in error types from pre-therapy (T3) to post-therapy (T4). NR=no response; SE=semantic error; PH=phonemic error; UNR=unrelated error; VIS=visual error; MIX=mixed error.

5.1.2.4 Therapy effects: Spoken Picture Description

Performance on the SPD task at each of the three time points (T2, T3 and T4) is displayed in Figure 5.6. A repeated measures ANOVA was employed to investigate improvements and this analysis had three factors, each with two levels: time (pre-therapy (T2-T3) and therapy (T3-T4); trained (trained (TR) and untrained (UN)); and tested (tested (T) and untested (UT)). Trained and tested items were those which were trained in the therapy and tested in the WRT; untrained but tested items were those which were untrained in the therapy but tested in the WRT; trained but untested items were words which were trained in the therapy but not tested in the WRT (see 'core lexical items'); and untrained and untested items were words which were not trained in the therapy or tested in the WRT. No significant interactions were found but there was an overall effect of time, $F(1, 26)=4.82$, $p=0.04$.

Due to this result, further analyses (a paired sample t-test and a Wilcoxon signed ranks test) were completed to examine change in the number of unique information

carrying words (ICWs) produced from pre-therapy to post-therapy (all item types), as well as the number of all trained items (both tested and untested in the WRT) (Figure 5.7). This revealed significant increases in the retrieval of ICWs and all trained items at the group level from the pre-therapy block to the therapy block (ICWs: *t*(26)=2.2, p=0.04; pre-therapy M=23.67, SD=15.44; post-therapy M=28.19, SD=15.88; all trained items: Z=2.32, p=0.02; pre-therapy M=6.41, SD=5.18; post-therapy: M=8.59, SD=5.77).

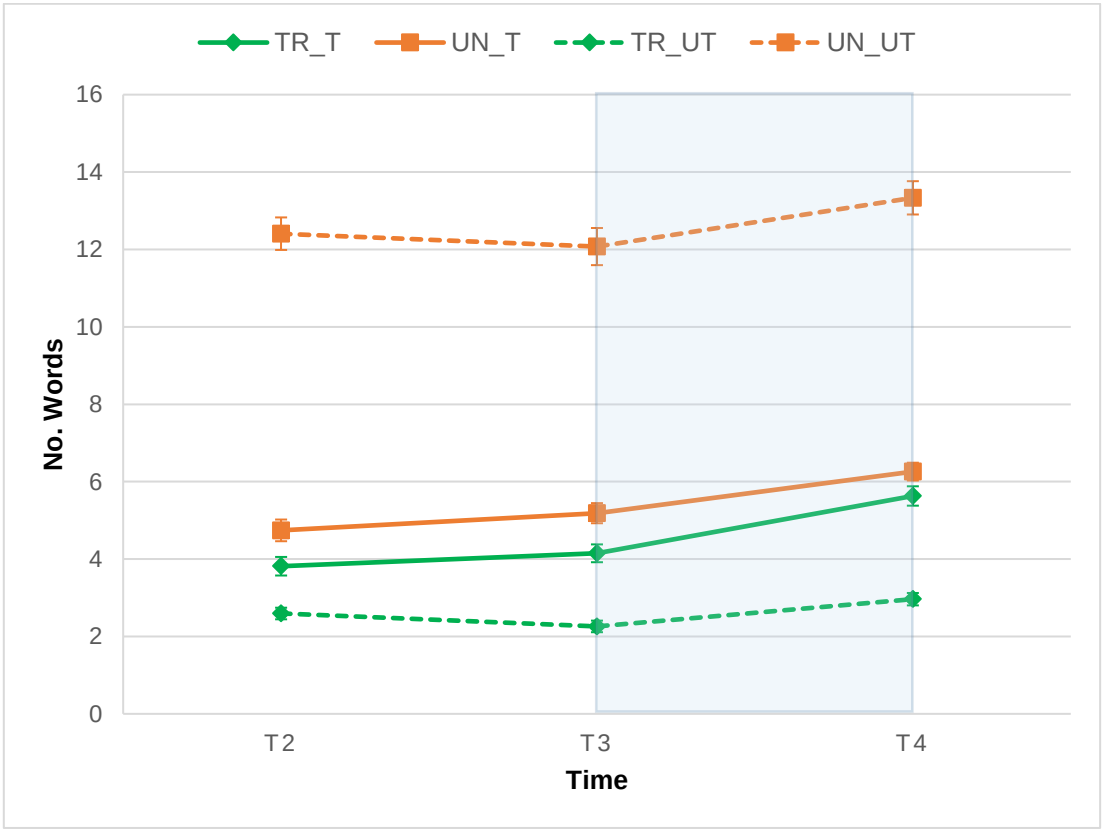


Figure 5.6 Performance (number of words produced) on the Spoken Picture Description task at T2, T3 and T4. TR_T=trained and tested items; UN_T=untrained but tested items; TR_UT=trained but untested items; UN_UT=untrained and untested items. The blue shaded area denotes the six week therapy block. Error bars are within-subject standard error of the mean.

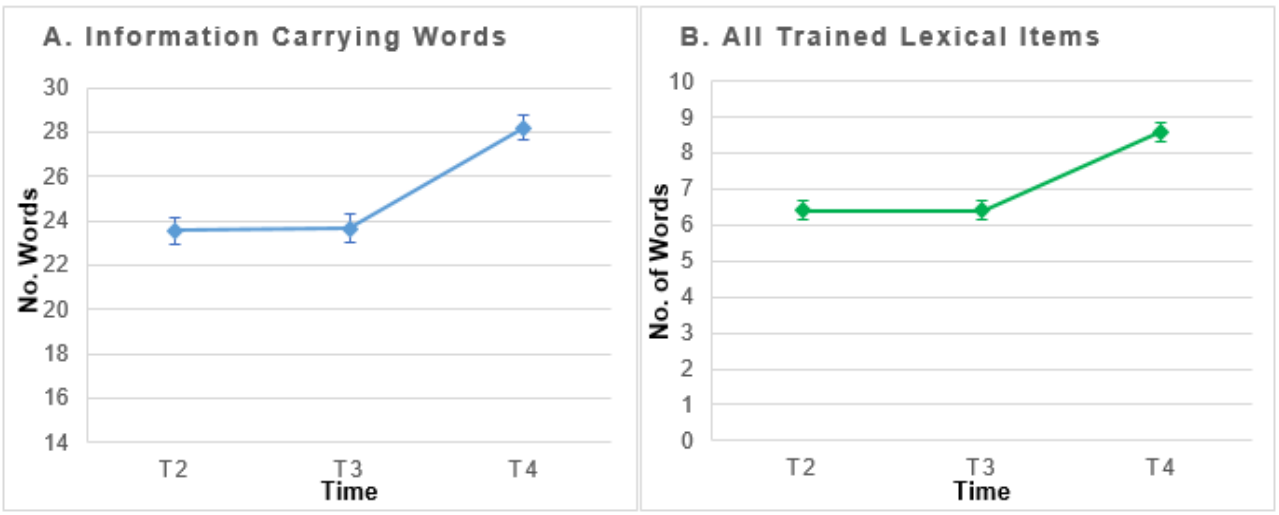


Figure 5.7 Performance (number of words produced) on the Spoken Picture Description task at T2, T3 and T4 for ICWs (a) and all trained items (b). Error bars are within-subject standard error of the mean.

5.1.2.5 Therapy effects: Self-reported outcomes

The DVAMS were completed at pre-therapy (T3) and post-therapy (T4) to assess participants' mood. P5 did not complete the questionnaire at T4 so results are reported for 26 participants. Although 16 participants reported an improvement in mood on the DVAMS following therapy, and the overall group mean increased from 71.2 (SD=16.2) to 74.54 (SD=14.68), a Wilcoxon signed ranks test showed this change was not significant, Z=-1.38, p=0.17.

5.1.2.6 Explaining response to treatment

There was variability across participants in the improvement of trained lexical items on both the WRT and the SPD task (for the SPD task, trained lexical items include (in this section) 'core lexical items'), and change on the two tasks was not significantly correlated, $r(27)=0.14$, p=0.5 (Figure 5.8). Improvement on each of these tasks following therapy varied between individual participants, with some making significant gains on the WRT but not the SPD (P7, P16, P18, P19, P22 and

P26) and others improving on the SPD but not the WRT (P5, P6, P8, P13 and P23) (Figure 5.9).

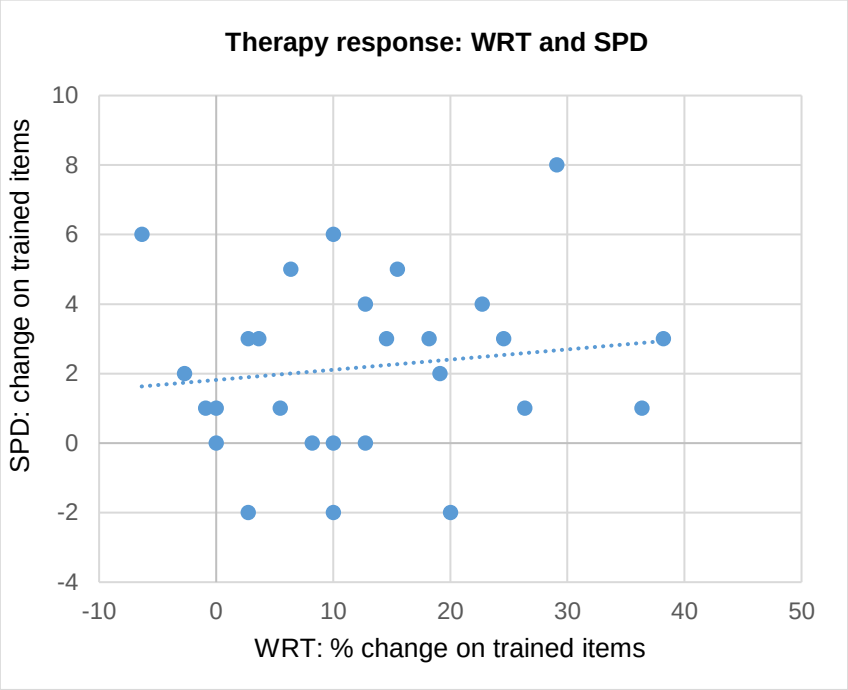


Figure 5.8 Scatterplot showing the correlation between change on trained items on the WRT and the SPD.

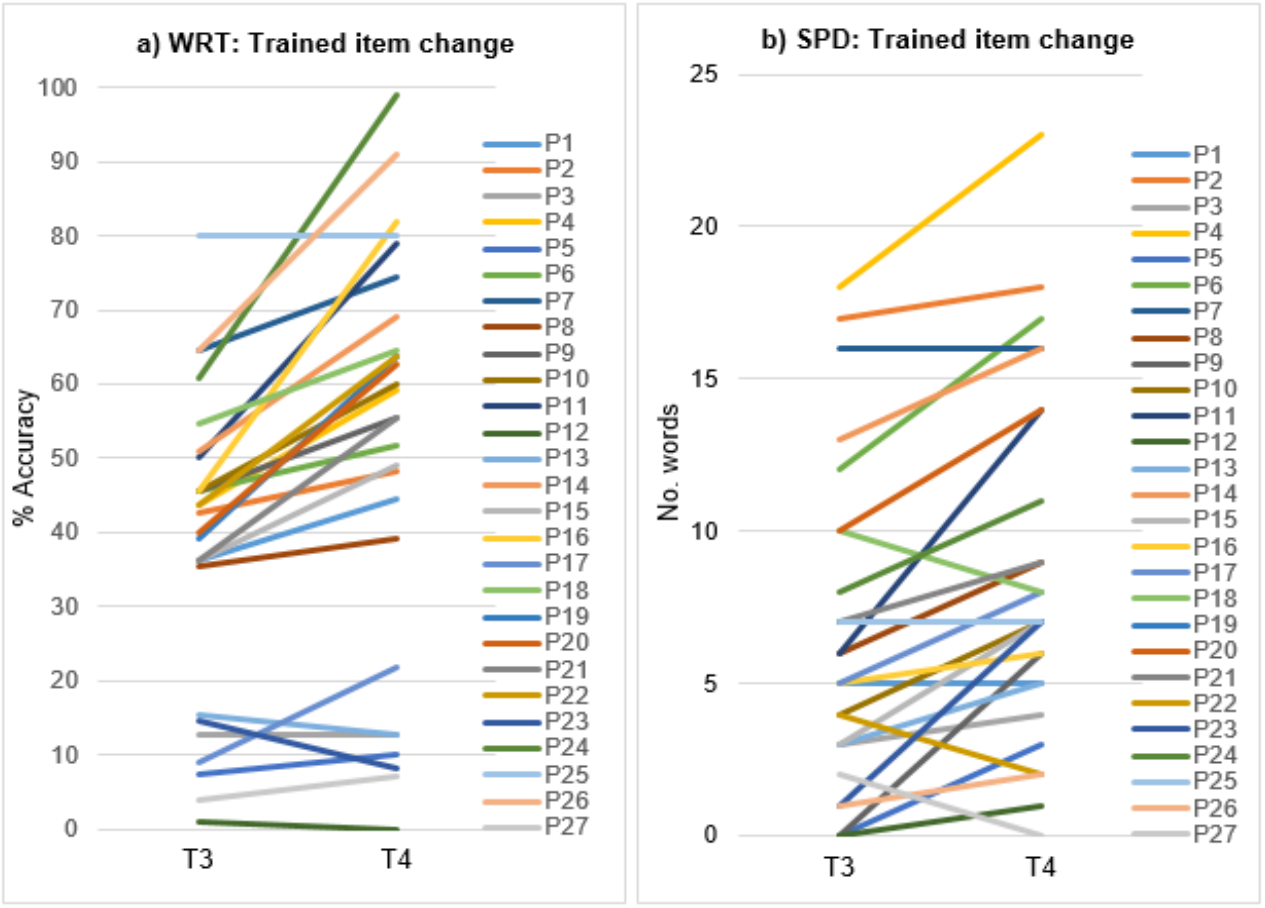


Figure 5.9 Line graphs showing change on the WRT and the SPD from T3 to T4. (a) % change on trained items on the WRT; (b) raw change on trained items on the SPD).

As there was no significant correlation between WRT and SPD change, simple correlations of therapy response were completed separately for each task. This analysis ascertained whether there were any relationships between improvement (T3-T4) and key explanatory variables (baseline severity on the WRT, dose and exposure). Results revealed that baseline severity, dose and exposure were not correlated with therapy response on either the WRT (baseline severity: $r(27)=0.13$, $p=0.5$; dose: $r(27)=-0.37$, $p=0.06$; exposure: $r(27)=0.18$, $p=0.36$) or the SPD (baseline severity: $r(27)=0.06$, $p=0.77$; dose: $r(27)=-0.01$, $p=0.97$; exposure: $r(27)=-0.03$, $p=0.89$) (Figure 5.10).

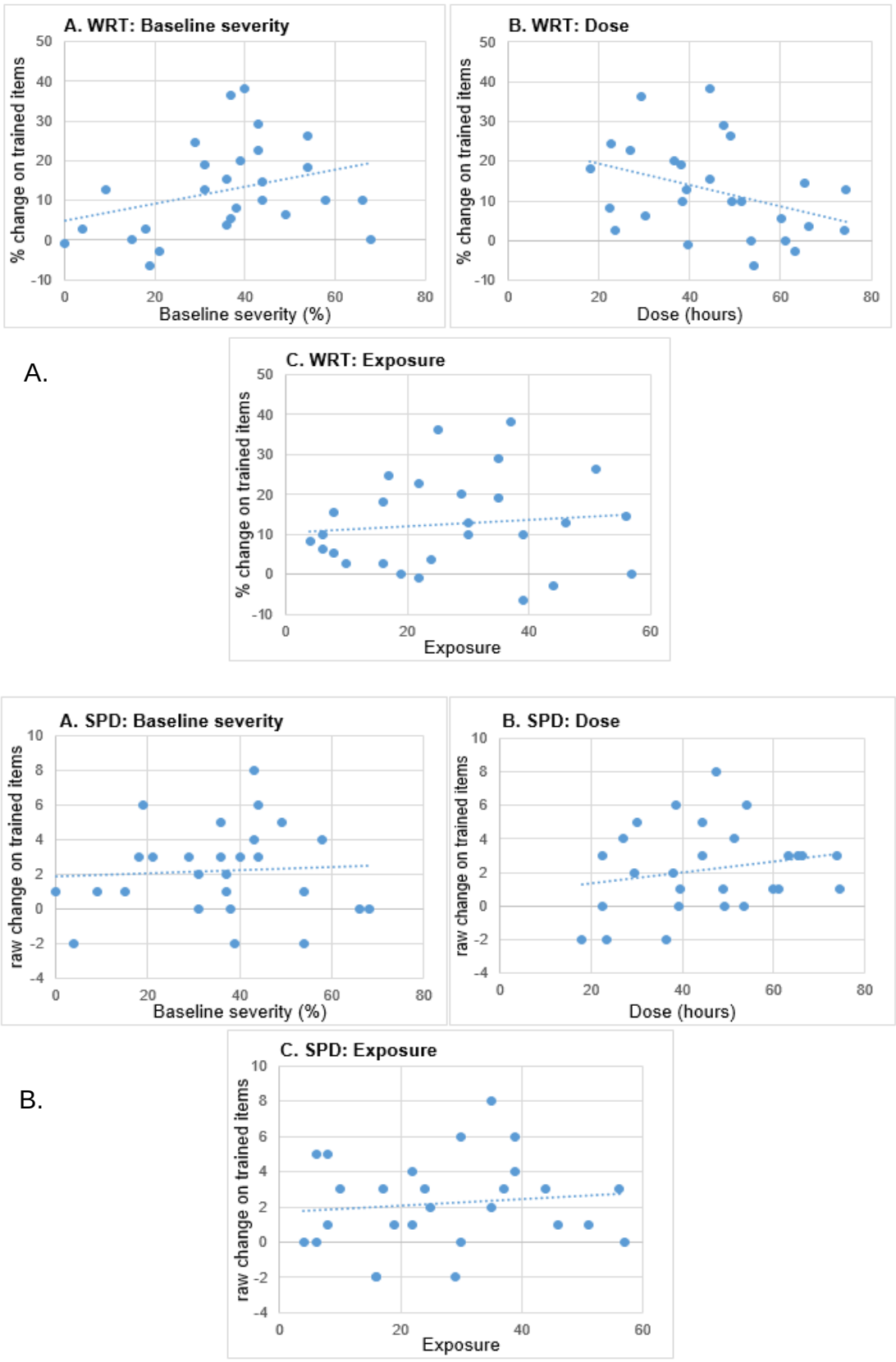


Figure 5.10 Scatterplots showing correlations between dependent variables. (A. % change on trained items on the WRT; B. raw change on trained items on the SPD; with possible explanatory variables (A: baseline performance; B dose; C: exposure).

5.1.2.7 Automatic Linear Modelling

As simple correlations could not account for the variability in individual response to therapy, ALM was used to assess if a combination of behavioural and demographic variables could explain some of the variance. Two separate analyses were conducted, each with a different dependent variable: analysis 1 examined percentage change on trained lexical items from T3 to T4 on the WRT; whereas analysis 2 investigated raw change on trained lexical items from T3 to T4 on the SPD task. Both models included 15 predictor variables, which were: dose, age, time since stroke (TSS), baseline performance on the WRT, the PALPA, the DVAMS and the CAT (subtests included: spoken picture description, naming objects, repetition, reading words, comprehension of spoken words and sentences, comprehension of written words and sentences, and digit span). To allow for inclusion of all 27 participants in the modelling, the cognitive assessments that were not completed by all participants were not used as predictor variables.

Analysis 1

The results for analysis 1, change on trained items on the WRT, identified a best model which included three predictor variables (Figure 5.11). The Akaike information criterion for this explanatory model was 83.1 and the adjusted R^2 was 0.897, suggesting that the model accounts for 89.7% of the variance in therapy response for the WRT. The significant variables in order of importance were: naming objects (β=-12.99, p=>0.001); baseline performance on the WRT (β=12.04, p=>0.001); and comprehension of spoken sentences (β=15.38, p=>0.001). The naming objects variable showed a negative correlation with improvement, suggesting those who performed poorer on this test improved the greatest on the WRT. Interestingly,

although scores on naming objects and scores on the WRT at baseline were highly correlated (see section 4.6.1 in Methods), WRT performance at baseline was positively correlated with change on the WRT from T3 to T4. Performance on the comprehension of spoken sentences also showed a positive correlation, indicating those who had greater scores on these two measures, had a greater response to therapy. The model did identify one participant who had an outlying dependent variable score (change in trained lexical items on the WRT), however, the Cook's distance was less than one (0.36), so this subject was not removed from the analysis.

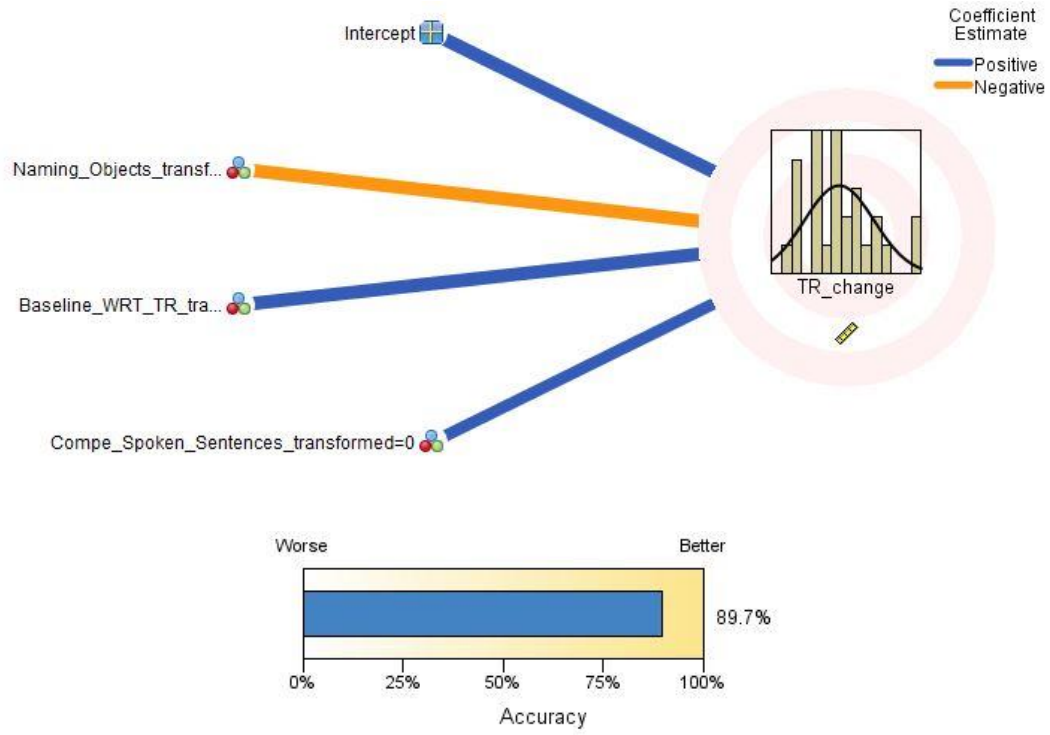


Figure 5.11 Automatic Linear Modelling: Change on the WRT. Variables are displayed in order of significance. Blue lines = positive correlations; orange line = negative correlation. The horizontal blue bar at the bottom represents model accuracy (adjusted R^2).

Analysis 2

The best model identified in analysis 2, change in trained items on the SPD task, is shown below in Figure 5.12. The Akaike information criterion for this explanatory

model was 30.38 and the adjusted R^2 was 0.814, suggesting that the model accounts for 81.4% of the variance in therapy response for the SPD task. The model included two significant predictor variables which were the PALPA (β =-3.14, p=0.001) and naming objects (β =5.83, p=0.001). Scores on the PALPA showed a negative correlation with change for trained items on the SPD, suggesting those who had worse non-word repetition at baseline were able to retrieve more trained items post-therapy. Conversely, to both the PALPA and the results from analysis 1, naming objects had a positive correlation with therapy response on the SPD task, indicating those who had higher scores on this subtest improved the greatest at post-therapy.

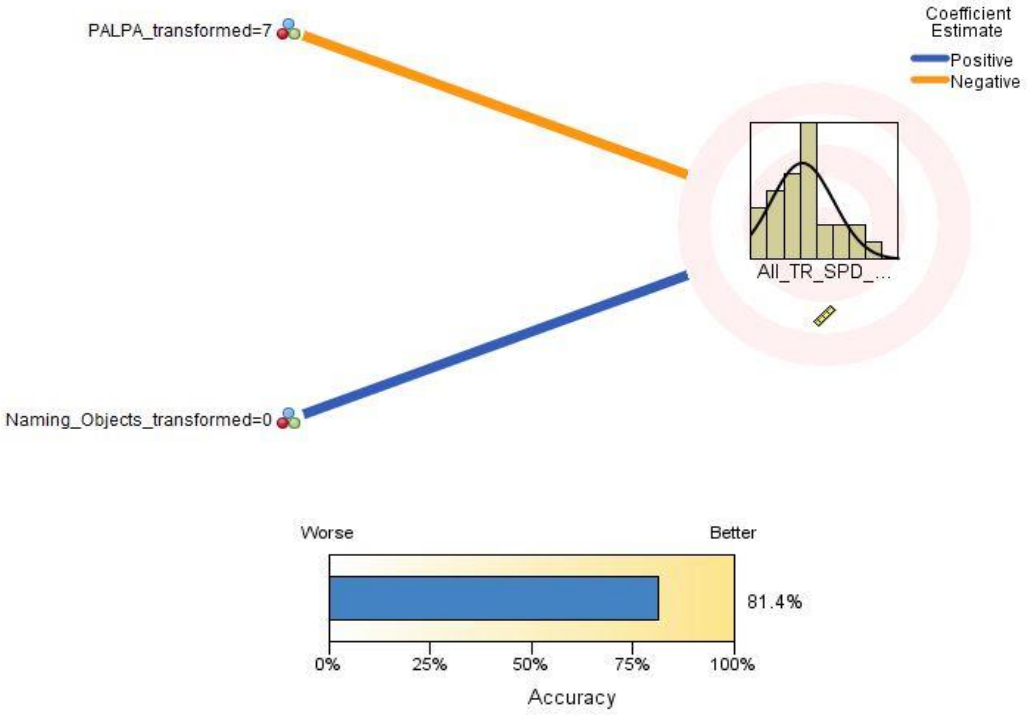


Figure 5.12 Automatic Linear Modelling: Change on the SPD. Variables are displayed in order of significance. Blue line = positive correlation; orange line = negative correlation. The horizontal blue bar at the bottom represents model accuracy (adjusted R^2).

5.1.3 Discussion

5.1.3.1 Summary of results

Six weeks of iTalkBetter significantly improved the retrieval of trained lexical items in participants with chronic post-stroke aphasia, and this was mirrored in a reduction of all error types from pre- to post-therapy. On average, participants improved by 13% and these gains were maintained 12 weeks following the cessation of therapy. No similar group level effects were found for untrained lexical items, supporting Hypothesis 1; that learning would be item-specific. Although there were small improvements in untrained lexical items, this was not significant and was likely due to the learning element of the 'fly-by' in the main outcome measure. Furthermore, in support of Hypothesis 2, participants made significant gains in the retrieval of both trained concrete and trained abstract words, suggesting the 'concreteness' of the word did not affect therapy response.

Participants also improved their retrieval of trained items in the spoken picture description task, demonstrating a generalisation of learning effects from the single word naming therapy to a less constrained task, supporting Hypothesis 3. Interestingly, a significant improvement was also found from pre-therapy to post-therapy in the number of unique information carrying words produced, suggesting an overall improvement in word retrieval in this task which was not specific to trained items. Although simple correlations could not explain response to therapy (for either the WRT or SPD task), a combination of behavioural baseline variables (but not demographic variables) did account for a large amount of the variance in the results (89.7% and 81.4%, respectively), supporting Hypothesis 4.

5.1.3.2 Comparison with previous findings

The iTalkbetter therapy utilised a hierarchical, error-reducing, phonological cueing treatment approach to improve word retrieval. Previous studies which have employed this type of the therapy technique have found similar positive outcomes in naming at the group level for PWA (Conroy, Sage & Lambon Ralph, 2009b; Conroy & Scowcroft, 2012; Nardo et al, 2017). The current results are also supportive of studies indicating the effectiveness of delivering aphasia therapy digitally (Brady et al, 2016; Zheng, Taylor & Lynch, 2016; Des Roches & Kiran, 2017). Specifically, as a computerised self-led therapy (CSLT), the iTalkBetter study provides further evidence that this type of cost-effective, high dose, impairment-based intervention is efficacious in improving language outcomes. Therefore, the present findings also align with those from Palmer and colleagues (2019) Big CACTUS randomised control study, which utilised the StepByStep© programme (Palmer et al, 2012), targeting the production of 100 personally relevant words with PWA over a six-month period. As in iTalkBetter, the 83 participants who received the therapy (were part of the experimental arm) made significant improvements in the retrieval of trained items (raw change: 16%). Furthermore, in iTalkBetter participants were able to independently use the app with minimal to no help from a communication partner or the research team, and age did not seem to impact the ability to engage with and use the therapy, in line with previous research (Munsell et al, 2020).

5.1.3.3 Generalisation

Following iTalkbetter therapy, participants improved in their production of trained items only, with no significant improvements in untrained items at the group level. In Hypothesis 1, I predicted that learning would be item-specific due to the limited

within-level generalisation reported across impairment-based aphasia therapy studies (Webster, Whitworth & Morris, 2015; Wisenburg & Mahoney, 2009; Raymer et al, 2008; Nickels, 2002). Although numerically there were improvements in untrained items over the iTalkbetter therapy period, this change was not different from the change observed over the pre-therapy block. Therefore, improvements in untrained items were likely due to the learning effects of the 'fly by' in the main outcome measure (the WRT). It may be, however, that there were individual differences in generalisation to untrained items, perhaps due to factors such as the level and severity of participants' lexical impairment, as proposed by previous research (Best et al, 2013; Holland, John & Woollams, 2018). As an analysis examining change for trained and untrained items at an individual level was not undertaken in this study, future work which investigates this would provide insight into whether within-level generalisation occurred for some participants and, if it did, why this might have been.

Importantly, although no significant within-level generalisation was found after six weeks of iTalkBetter, there was a significant improvement at the group level on the spoken picture description task, demonstrating across-level generalisation and supporting Hypothesis 3. Although the SPD was a less constrained task, in comparison to the therapy, with multiple target items within a scene, the stimuli still consisted of static pictures. Therefore, the across-level generalisation observed in SPD is an instance of 'near transfer' and is aligned with previous research suggesting that generalisation is more likely to occur if the skills required to complete a task are comparable to those in which the learning occurred (Subedi, 2004; Conroy, Sage and Lambon Ralph, 2009c). For example, in Conroy, Sage and Lambon Ralph's study, greater improvements were observed in picture supported

narratives as opposed to unsupported narratives following single word therapy in a group of seven PWA.

Due to the across-level generalisation observed, important next steps for the iTalkBetter therapy programme would be to examine improvements across the whole communication continuum, from supported and unsupported narratives, to conversation. As many PWA continue to face difficulties in conversation, despite gains made in therapy, these investigations are important to identify the steps required to promote across-level generalisation to further novel contexts (Carragher, Conroy, Sage & Wilkinson, 2012).

Another important point to note is that, in addition to a significant improvement in the retrieval of trained items in the SPD, there was also a significant improvement in the number of unique information carrying words produced at post-therapy. This suggests improvements on the SPD were not restricted to trained items only and, possibly, there was within-level generalisation from trained to untrained items on this task. Alternatively, changes in performance at post-therapy on the SPD may have been influenced by changes in more general cognitive abilities, such as attention, executive control and memory, which occurred due to engagement with the iTalkBetter therapy. Future analysis which investigates change at an individual level for trained and untrained items on the SPD (as discussed above for the WRT) would provide information on whether there was within-level generalisation on this task. However, as an in-depth post-therapy evaluation of cognition was not completed, it would impossible to discount the possibility that changes in cognitive factors were driving this result.

5.1.3.4 Concrete and abstract items

A novel feature of iTalkBetter, in comparison to the majority of impairment-based aphasia therapies, is that (in addition to concrete highly imageable words) it also included abstract words with low imageability. Previous research has indicated the need to include more abstract words in therapy due to the frequency of these word types in the English language, and the importance of these words for expressing more complex needs and emotions (Renvall, Nickels & Davidson, 2013; Worrall et al, 2011). Additionally, a single case study by Renvall and Nickels (2019) demonstrated that a traditional therapy technique (repetition in the presence of a picture) was effective in improving the retrieval of abstract adjectives. As there was also a significant improvement in both trained concrete and trained abstract words following iTalkBetter therapy, the current study provides further evidence (with a much larger participant group) that traditional therapy approaches may be appropriate for targeting more abstract items, supporting Hypothesis 2.

A factor which was not addressed in this analysis was whether there was a difference between abstract and concrete words in terms of how many repetitions of, or exposures to, each item were required to improve the retrieval of that item. It has been established that PWA exhibit a 'concreteness' effect, in which concrete, highly imageable, words are easier to retrieve than abstract words and require less cueing (Walker & Hulme, 1999; Conroy, Snell, Sage & Lambon Ralph, 2012; Nickels & Howard, 1995; Kiran, Sandberg & Abbott, 2009). Therefore, in iTalkBetter, future work should aim to capture changes over the course of therapy to examine the relationship between improvements in concrete and abstract words and the number of exposures. Moreover, in the present analysis, the comparison of concrete and abstract items was completed based on a median split of the words using

concreteness ratings. This method may not have been the most appropriate way to investigate changes in the two different word types as concreteness is really along a continuum. Due to this, a further consideration for this work would be to examine changes in retrieval along the concreteness continuum.

5.1.3.5 Dose and intensity of therapy

Over the course of the six-week block of iTalkBetter therapy, participants completed a relatively large dose of therapy, with an average of 45 hours. This amount of therapy is within the recommended dosage required to improve language functions as proposed by previous research (20-100 hours) (Brady et al, 2022; Brady et al, 2016; Bhogal et al, 2003). Overall, participants tolerated this high therapy dose as no individuals dropped out due to not being able to engage with the treatment programme, in line with a study completed by Breitenstein and colleagues (2017). However, the research team did provide regular encouragement to the participants, as in those studies by Palmer (2019), Fleming (2020), Woodhead (2017), and colleagues. It is likely this contact with the research team contributed to the high dose completed as emails and phone calls were often used to motivate those who were falling behind in the intervention. Despite the encouragement provided, six participants achieved less than half of the target dose (<30 hours) due to reasons such as time constraints, illness, fatigue, boredom, low motivation and difficulties with the task. Therefore, as suggested in Brady and colleagues' Cochrane review (2016), this type of therapy may not be suitable for all individuals. On the whole, however, high dose CSLT was achievable for the majority of participants.

In the aphasia literature, dose of therapy (number of hours achieved) and intensity of therapy (frequency of treatment) is poorly defined, with many studies confounding

the two, and others reporting different intensity schedules which relate to either ‘high intensity’ or ‘low intensity’ practice. In the present study, a dose of 45 hours over six weeks equates to roughly seven and a half hours of therapy a week, which is considered both a high (Stahl et al, 2018) and low (Dignam et al, 2015) intensity schedule, depending on individual viewpoints. It is important to note that the iTalkBetter clinical trial did not aim to manipulate dose and intensity and did not set out to assess the relationship between dose and therapy response, or compare intensity schedules. However, simple correlations were undertaken to examine the linear relationship between dose, intensity and therapy response but none were found. It is, therefore, unclear what the optimal dose or intensity schedule is at the group and individual level. Future work that systematically manipulates the dose and intensity of the iTalkBetter therapy (for example, through an adaptive trial design), while also examining individual factors (such as aphasia severity, time since stroke and age), would provide clarity on how much and how frequent therapy should be for each PWA (Brady et al, 2022; Doogan et al, 2018; Bhatt & Mehta, 2016).

5.1.3.6 Explaining response to therapy

Participants varied considerably in terms of how well they responded to iTalkBetter therapy, on both the WRT and the SPD. In the explanatory modelling completed, object naming (CAT), baseline performance on the WRT and comprehension of spoken sentences (CAT) were found to explain a large amount of the variance in change on the WRT (89.7%). Whereas for change on the SPD, non-word repetition (PALPA) and naming objects were found to explain a large amount of the variance (81.4%). As object naming was significant for both models, it suggests this factor has an important explanatory value for overall therapy response. However, for the WRT the relationship between therapy response and object naming was negative (those

who were more severe at baseline improved the most), and for the SPD the relationship was positive (those who were less severe at baseline improved the most). These differential findings reflect the inconsistencies in the literature between baseline severity and treatment outcomes, with some reporting those with milder impairments do better (Lambon Ralph, Snell, Fillingham Conroy and Sage, 2010; Kiran, 2016), and others finding those with more severe deficits improve the most (Laska et al, 2001; Robey, 1998). In iTalkBetter it seems that baseline naming severity has contrasting effects on treatment response, with those who are milder improving more at the single word level, and those who are more severe generalising their learning to a less constrained task.

Finding relationships between baseline factors and therapy outcomes is clinically important as PWA devote considerable time to treatment programmes and identifying who responds best to a particular therapy enables the provision of suitable interventions at an individual level. Although the current models explain a large amount of the variance, future modelling analysis which also includes structural imaging data may further contribute to understanding the variability between participants in response to therapy (Aguilar et al, 2018). Additionally, the analysis completed in the present study is explanatory only and the models do not provide predictive information on who is likely to improve following iTalkBetter. Predictive analysis may identify that the explanatory models over-fit the data in the present research sample. For example, in Aguilar and colleagues' study, the researchers conducted an out-of-sample analysis to assess the predictive power of their models and found that, although the combined model explained 94% of the variability in treatment response, it had considerably less predictive power (23%).

5.1.3.7 Limitations

Following six weeks of iTalkBetter therapy, participants significantly improved their retrieval of trained items and this was maintained three months following the end intervention. The primary limitation of this study, however, is the differing number of items in the therapy corpus for the first seven participants (n=890), in comparison to the following 22 participants (n=220). The reduction in the number of items provided within iTalkBetter during the study was due to the time it took to work through the entire word corpus, leading to limited exposures to each trained item over the six weeks. Although reducing the word list greatly increased the number of exposures and simple linear regressions did not find an association between treatment response and exposure count, this is likely to have affected therapy outcomes. However, it is probable that having a larger therapy corpus weakened the overall group level result and, as iTalkBetter did significantly improve word retrieval, this further strengthens the current results.

An additional limitation is that, due to the COVID-19 pandemic, seven participants completed the study entirely remotely with no face-to-face contact with the research team. Although this provides evidence that iTalkBetter is suitable as a self-led therapy which can be completed independently at home, there may have been differences between those who attended testing sessions at UCL and those who did not, for example, in factors such as motivation and engagement. Furthermore, many of the cognitive assessments were not able to be completed remotely so could not be included in the explanatory modelling. As previous studies have highlighted the importance of baseline cognitive factors in explaining response to therapy (Yeung & Law, 2010; Lambon Ralph et al, 2010; Dignam et al, 2017), the inclusion of these

tests within the models may have provided more detailed information regarding the variability in treatment outcomes.

5.1.3.8 Summary and future considerations

The current study adds to the research evidence for aphasia therapy by demonstrating the effectiveness of a computerised, self-led therapy in improving the retrieval of trained words. On average, participants improved their retrieval of trained items by 13% and these treatment gains were maintained three months following the cessation of therapy. As iTalkBetter is an independent, home-based therapy programme, these results have the potential to be easily translated into real-word clinical practice. The iTalkBetter the app will soon be released online in a Phase III 'roll-out' study where any person with aphasia will be able to download and use the therapy, and both the WRT and the SPD will be implemented into the app to enable further investigations of therapy response on a larger scale.

In the roll-out, phase III study, a number of factors could be manipulated to investigate how iTalkBetter could be optimised to maximise therapy outcomes, for example, by using A/B testing methods in which two versions of the therapy are compared against one another. These may include providing the written form of the word due to research indicating the benefits of orthographic cues (Sze et al, 2021); increasingly requiring quicker word retrieval to potentially facilitate improved generalisation to connected speech (Conroy et al, 2018); enabling participants to choose words practiced in therapy so the items are functional and personally relevant (Webster, Whitworth & Morris, 2015); and manipulating dose and intensity to examine ideal schedules at an individual level (Brady et al, 2022). Future work could also examine an adaptive form of iTalkBetter in which words are removed from

the therapy content when they are considered ‘learnt’ and new words are introduced. Additionally, user feedback could be collected to identify areas in which the app could be developed at a usability level to improve engagement with the therapy.

5.2 Chapter 2: Investigating structural brain adaptation in response to iTalkBetter therapy.

Aim: To investigate whether iTalkBetter therapy induces changes in the structural architecture of the brain in persons with chronic aphasia.

Hypotheses:

(1) iTalkBetter therapy will induce increases in volume in the grey and/ or white matter of the language and/ or cognitive networks.

(2) Increases in volume in the grey and/ or white matter will correlate with percentage change on trained items on the WRT from pre- to post-therapy.

(3) Dose of therapy (number of hours completed) will correlate with increases in the volume of grey and/ or white matter.

5.2.1 Experimental procedures

5.2.1.1 Participants

In the following analyses, data from a subset of 17 participants, who were medically suitable and available for MRI scanning (dependent on the COVID-19 situation), are presented. This subset of participants were representative of the entire cohort in baseline performance on the WRT, response to treatment (percentage change for trained items on the WRT from pre-therapy (T3) to post-therapy (T4)), dose of therapy, exposure to trained items, age and time since stroke (TSS) (Table 5.3).

Group	Baseline WRT (%)	WRT change (%)	Dose (hours)	Exposure	Age	TSS
Whole cohort	32 (16)	13 (12)	45 (16)	27 (16)	62 (9)	83 (67)
MRI cohort	32 (16)	14 (13)	46 (16)	27 (16)	62 (7)	68 (73)

Table 5.3 Behavioural and demographic variables for all participants and for the subset of participants who took part in MRI scanning. Numbers in brackets are standard deviations.

In Figure 5.13 is the lesion overlap map, displaying the distribution of participants' lesions across the brain. All participants had LH lesions, with the majority of participants having extensive and widespread damage, encompassing the temporal, parietal and frontal lobes. Maximal overlap was in the white matter tracts of the corona radiata and the corpus callosum, and in the grey matter of the supramarginal gyrus. Twelve participants had lesions which extended into Broca's area (inferior frontal gyri: pars triangularis and pars opercularis) and twelve had damage to Wernicke's area, in the superior temporal gyrus. No participants had lesions in the RH.

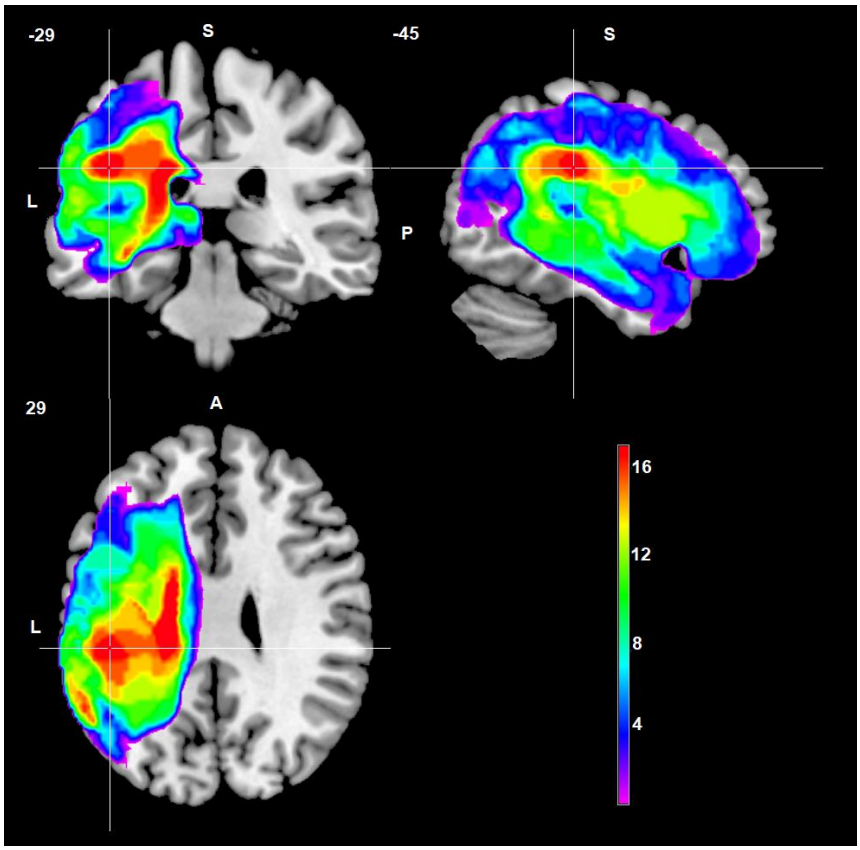


Figure 5.13 Lesion overlap map displaying the distribution of participants' lesions. S = superior; P = posterior; A = anterior; L = left. The colour bar represents the number of participants with a lesion in a voxel, from 1 (purple) to 17 (red). Numbers represent MNI coordinates.

5.2.1.2 Design

The scanning protocol is outlined in the Methods. A repeated measures design was used, with three scanning time points: two pre-therapy (T2 and T3) and one post-therapy (T4).

5.2.1.3 Data pre-processing

Following the prior pre-processing steps, described in the Methods, the difference images produced were smoothed using an isotropic kernel of 8mm full-width half-maximum (FWHM).

5.2.1.4 Statistical analyses

The difference images represented within-subject change in volume of GM and WM over the six-week therapy block (T3-T4) in comparison to the six-week pre-therapy block (T2-T3). The 17 smoothed GM and WM difference images were entered into three separate simple linear regression models in SPM12.

Analysis 1: To investigate change in volume of GM and WM as an overall effect of partaking in the therapy program, the 17 contrast images were entered into a second-level group analysis with no covariates.

Analysis 2: To identify if there were changes in volume which correlated with response to therapy, percentage change on trained items from pre-therapy (T3) to post-therapy (T4) on the WRT was entered as the dependent variable in the second model (Figure 5.14a).

Analysis 3: Dose (number of hours of therapy completed) was entered as the dependent variable in the third model to ascertain if changes in volume correlated with the amount of therapy participants achieved (Figure 5.14b).

For Analyses 2 and 3, exposure was entered as a covariate of no interest due to differences in exposure to trained items across participants (see Methods and Chapter 1). In the pre-processing steps, between subject-effects (baseline behavioural and demographic variables) were accounted for in the within-subject design, so no further regressors were entered into the models. The statistical voxel-level threshold was set at $p<0.001$ uncorrected, and to correct for multiple comparisons, significant clusters are reported at the Family Wise Error (FWE) $p<0.05$ threshold.

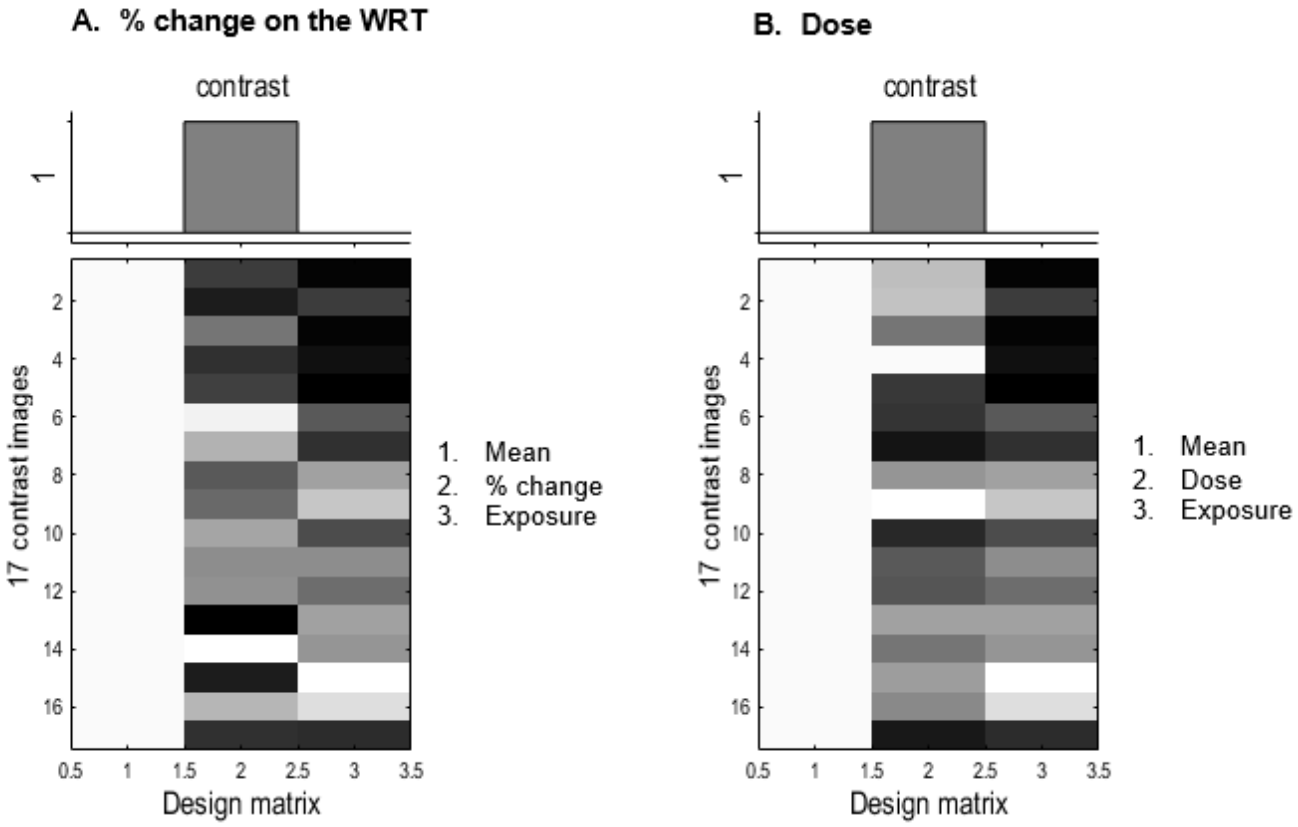


Figure 5.14 Design matrices used for GM and WM linear regression models. One regressor of interest was included in each of the models for Analyses 2 and 3: (A.) change on WRT trained items (%); (B.) dose (number of hours completed). Both models also included exposure as a covariate of no interest.

5.2.2 Results

5.2.2.1 Analysis 1

In Analysis 1, change in volume in GM and WM was investigated from pre-therapy to post-therapy with no added regressors in the model. This analysis identified significant clusters which increased in volume at post-therapy in both GM and WM, supporting the first Hypothesis, and are displayed in Figure 5.15 and Table 5.4. In the GM, a significant cluster was found in the LH in the anterior prefrontal cortex (aPFC) (BA10) and in the RH in the superior temporal gyrus (STG). The peak voxel in the RH was in the planum temporale (PT, BA22), the region contralateral to Wernicke's area in the LH, with a further peak in the cluster extending into the middle temporal gyrus (MTG, BA21). In the WM, a significant cluster which increased in volume at post-therapy was also found in the RH, in the white matter tracts underlying BA22 and BA21. The boxplots in Figure 5.16 display the change in volume in each of these three clusters for individual participants. P3 showed a decrease in volume across all three clusters; P24 was the only participant who had a decrease in volume in the GM of the aPFC; and one participant, P19, showed a decrease in volume in the WM of BA22. All other participants displayed increased brain tissue volume across all three clusters.

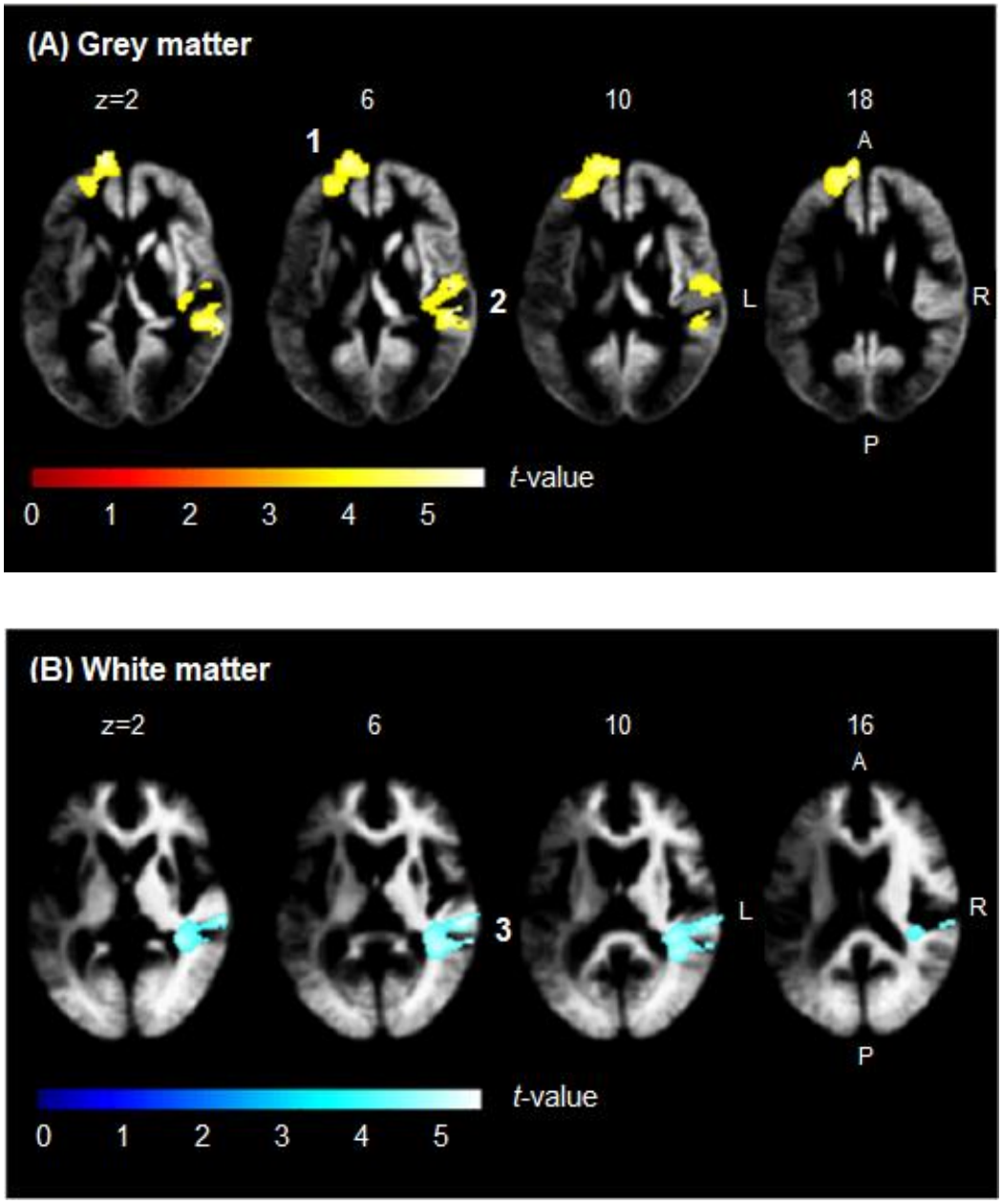


Figure 5.15 Significant clusters with increased volume at post-therapy. (A) Grey matter; (B) White matter. (1) GM: left aPFC; (2) GM: right PT; (3) WM: right PT. Brain templates are the average GM and WM images for the 17 participants. A = anterior; P = posterior; L= left; R = right. Colour bars represent t-values at that voxel. Numbers are MNI coordinates.

A. Grey matter					
Left hemisphere	***k***	***T***	***x***	***y***	***z***
Anterior prefrontal cortex (aPFC)	1356	5.68	-8	62	22
		5.54	-14	72	2
		5.24	-6	66	12
Right hemisphere	***k***	***T***	***x***	***y***	***z***
Planum temporale (PT)	957	5.92	56	-30	8
Middle temporal gyrus (MTG)		5.66	62	-36	2
Planum temporale (PT)		5.35	50	-34	4

B. White matter					
Right hemisphere	***k***	***T***	***x***	***y***	***z***
Planum temporale (PT)	982	5.05	60	-26	4
Middle temporal gyrus (MTG)		4.82	62	-40	6
Planum temporale (PT)		4.65	40	-46	8

Table 5.4 Significant clusters with increased volume at post-therapy. (A) Grey matter; (B) White matter. The numbers represent MNI coordinates of the first three peak voxels in the clusters; 'k' is cluster size; and 'T' is the t-value.

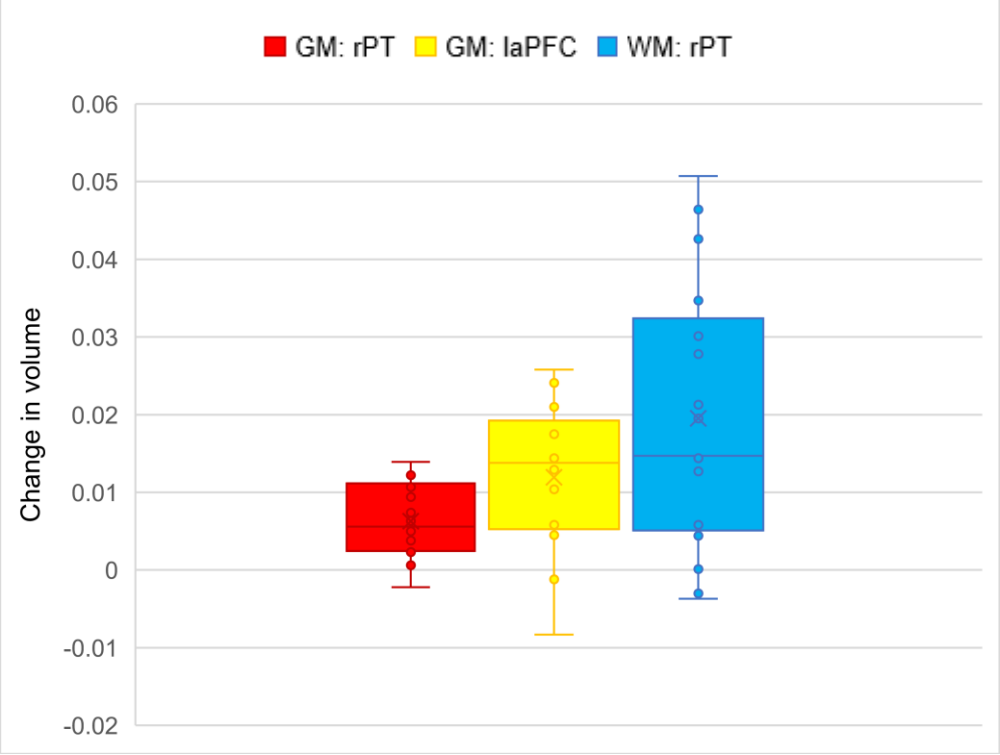


Figure 5.16 Boxplots showing individual change in volume in GM and WM. Red = GM of the right PT; Yellow = GM of the left aPFC; Blue = WM of the right PT. Volume values are eigenvalues from 5mm spheres around the peak voxel in each cluster.

5.2.2.2 Analyses 2 and 3

In analysis 2, whole brain MRI investigations did not identify any significant changes in volume from pre-therapy to post-therapy which covaried with response to therapy (change in trained items on the WRT). Additionally, no significant clusters were found when dose was used as the dependent variable in the model in analysis 3. These results do not support Hypotheses 2 and 3, indicating that volume changes in the GM and WM were not associated with response to therapy or time spent on therapy.

5.2.2.3 Post-hoc investigations

Following the increases in volume in GM and WM found in analysis 1, post-hoc investigations were undertaken to understand what structural changes were occurring in the brain over the pre-therapy and therapy period separately. To do this, the probabilistic change maps for GM and WM for pre-therapy and therapy blocks (detailed in step 4 of the pre-processing pipeline in the Methods) were spatially normalised into MNI space, smoothed at 8mm FWHM and entered into one sample t-tests in SPM12, looking at increases (0 1) and decreases (0 -1) in volume. This analysis was purely exploratory to understand the mechanisms of change so a lenient threshold of voxel-level significance was chosen ($p<0.01$). Due to the significant clusters identified in the frontal and temporal lobes, a small volume correction was completed using a temporal and frontal lobe mask from the WFU_PickAtlas (Maldjian, Laurienti, Kraft, & Burdette, 2003).

No significant areas of change were found at either the cluster or the peak level for GM or WM. Furthermore, no clusters were found that increased in volume over the pre-therapy block, or decreased in volume over the therapy block. In the GM,

however, there was a decrease in volume over the pre-therapy block in a cluster in the RH, with the peak voxel located in the primary auditory cortex (PAC) and further peaks in the PT (k=147) (Figure 5.17). Also, in GM, increases in volume were identified over the therapy block, with the largest cluster in the left aPFC (k=454). Similarly, in the WM, there were decreases in volume over the pre-therapy block and the largest cluster was in the right temporal lobe, with the peak voxel in the insula but also comprising the PAC and the PT (k=552). Over the therapy block, the largest cluster which increased in volume in the WM was in the left aPFC (k=227).

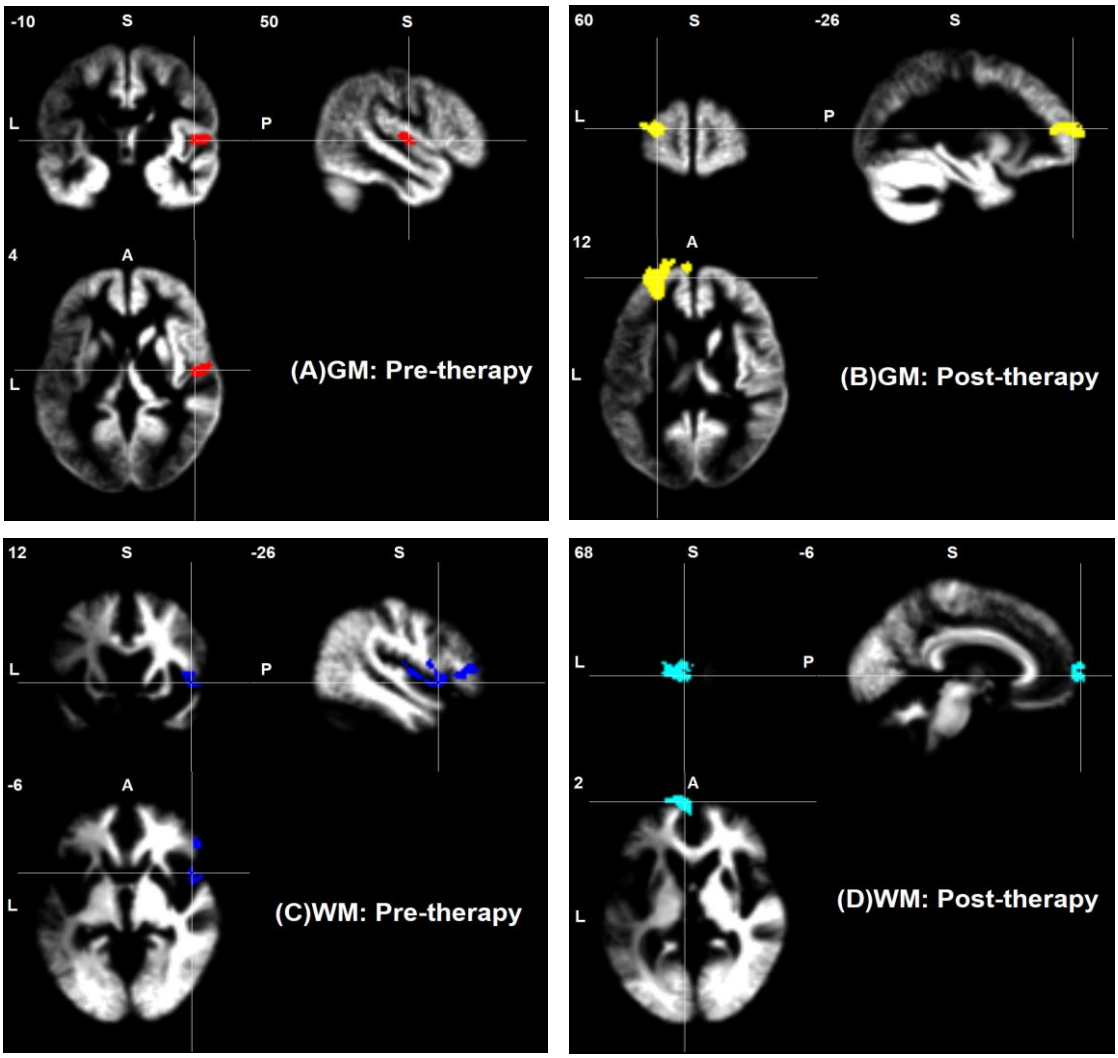


Figure 5.17 Largest clusters found in the GM and WM over the pre-therapy and therapy block separately. Brain templates are the average GM and WM images for the 17 participants. A = anterior; P = posterior; S = superior; L= left; R = right. Numbers are MNI coordinates.

5.2.3 Discussion

5.2.3.1 Summary of results

In this chapter, changes in the structural architecture of the brain were examined from pre-therapy to post-therapy in a subset of participants who took part in the iTalkBetter study. In Hypothesis 1, it was predicted that six weeks of therapy would induce increases in volume in the grey and/ or white matter within the language and/ or cognitive networks and the results support this hypothesis. In the GM, two significant clusters were identified: in RH in the planum temporale (BA22), the homologue region to Wernicke's area; and in the LH in the anterior prefrontal cortex (BA10). A significant cluster was also found in the WM tracts underlying BA22 in the RH. Post-hoc analyses revealed these changes may have been brought about by a cessation of natural atrophy in BA22 and hypertrophy in BA10 over the six weeks of iTalkBetter therapy.

It was also hypothesised that changes in GM and WM would correlate with response to therapy (change in trained items on the WRT) and dose of therapy (number of hours completed), however, this was not supported as no significant clusters were found which covaried with either of these dependent variables.

5.2.3.2 Comparison with previous findings

In longitudinal structural imaging studies with healthy individuals, experience-dependent plasticity has been observed in GM and WM in response to training in skills such as juggling and playing golf (Draganski et al, 2004; Bezzola, Merillat, Gaser & Jancke, 2011). In the language domain, structural changes have also been found following foreign language learning. For example, Mårtensson and colleagues (2012) examined neuroplasticity in a group of 14 interpreters who were completing

an intensive, three-month, foreign language training programme which involved learning 300-500 new words each week. In comparison to control subjects, the interpreters showed increased cortical thickness in the IFG, STG and middle frontal gyrus of the LH, and in the hippocampus of the RH.

Longitudinal studies which include structural imaging before and after aphasia therapy are rare. However, the ones that have been completed have identified changes in both GM and WM within intact LH regions and in language homologue areas in the RH (Schlaug, Marchina & Norton, 2009; Wan et al, 2014; Chang et al, 2021; Fleming et al, 2020). Three of these studies investigated changes in speech production, of which two examined changes in WM networks using diffusion tensor imaging (DTI) (Schlaug, Marchina & Norton, 2009; Wan et al, 2014), and one which detected increases in GM by employing diffusional kurtosis imaging (DKI) (Chang et al, 2021). Although the study by Fleming and colleagues investigated neuroplasticity in relation to speech comprehension, the authors utilised the same longitudinal VBM methodology following T1 structural scanning as used in this study. At the time of their research, this methodology was a relatively new process which had not previously been employed to examine therapy-induced neural adaptation in a population of aphasic individuals. The findings of structural changes in biological plausible brain regions in the current study adds to the viability of longitudinal VBM as an analysis method in aphasia research. It is also the first study, to my knowledge, to utilise this method to examine neural adaption in response to anomia treatment in chronic post-stroke aphasia.

Unlike the studies by Fleming and Wan, however, changes in GM and WM were not related to therapy outcomes, or even dose of treatment. Although the within-subject design of this study took into account heterogeneity across participants for factors

such as age and time since stroke, the mass-univariate analysis employed treated the individuals as one homogenous sample. Across PWA, it is likely there are between-subject differences recovery, leading to inconsistent patterns which correlate with improved language skills. Indeed, previous fMRI research has shown vast differences in the specific brain areas found to undergo neuroplasticity in response to treatment, and it is still unclear whether intact regions in the LH, RH homologues or bilateral adaptations support recovery (Thompson, 2017). Additionally, others have found neural changes following therapy may differ depending on factors such as lesion site and location (Fridriksson et al, 2006; Vitali et al, 2007). Therefore, it is perhaps understandable that in the current study (which did not recruit participants based on lesion characteristics), no single significant cluster was identified, across the highly varied group, which correlated with therapy outcomes.

5.2.3.3 Left hemisphere contributions

In the LH, increases in GM were found in the aPFC (BA10), a part of the salience (cingulo-opercular) network (Dosenbach et al, 2008; Stockert et al, 2020). As discussed in the literature review, BA10 is one of the largest regions in the human brain, though its exact function is still unknown. Studies have discovered links between BA10 and multiple cognitive functions including, but not limited to, cognitive processing, memory, organising behaviour, attention, awareness of competence, and inhibition (Burgess & Wu, 2013; Burgess, Dumontheil & Gilbert, 2007; Gilbert et al, 2006; Fleming, Weil, Nagy, Dolan & Rees, 2010; de Zubicaray, Zelaya, Andrew, Williams & Bullmore, 2000; Seeley, 2019). In macaque monkeys, research suggests BA10 is connected to auditory and multisensory areas of the STS as well as the ventral region of the insula, via the extreme capsule (Petrides & Pandya, 2007).

Furthermore, activations in BA10 in fMRI studies have been found in almost all types of cognitive tasks, leading to the hypothesis that it is a hub for metacognition and may be an attentional gateway, enabling an individual to control attention between inner thought and the external world (Burgess & Wu, 2013).

Within the aphasia literature, changes in activation in the salience network in fMRI studies have been identified, with researchers concluding that recruitment of these brain regions involved in general cognitive processing may be supportive of recovery (Hartwigsen & Saur, 2019). In a 2020 study, Stockert and colleagues investigated neural adaptation from the acute to the chronic stage of recovery in a group of 34 PWA. In the chronic phase, both participants with frontal lesions and those with temporo-parietal lesions showed increased activation in preserved parts of the language network, in the central executive network (fronto-parietal network), and in the salience network including the insula, the dorsal anterior cingulate cortices (dACC), and the supplementary motor area. Additionally, for those with temporo-parietal damage, increased activity in the left insula correlated with greater improvements in language skills (comprehension and production).

Although the above study examined spontaneous recovery, fMRI research assessing neuroplasticity in response to anomia therapy has also found changes in the salience network. Following an error-reducing phonemic cueing therapy (which was similar to the iTalkBetter therapy programme), Nardo, Holland, Leff, Price and Crinion (2017) identified increased activation in areas of the salience network in the RH in a group of 18 PWA. These areas were associated with immediate improvements, and long-term maintenance (three months post-therapy), in single word naming, and included the anterior insula and the dACC. In another error-reducing phonological cueing therapy study with three participants, Fridriksson and

colleagues (2006) found that bilateral activity in the aPFC correlated with naming improvements post-treatment for the participant with the smallest lesion.

The current study adds to those aforementioned by providing evidence of structural adaption in the salience network following an error-reducing, phonemic cueing therapy. As in Fridriksson et al's study, the location of this change was in the aPFC, however, neuroplasticity here was found for 15 of the 17 participants, rather than just one (see Figure 5.16). Dissimilar to Fridriksson's results, one of the participants (P24, see Methods) who did not show increased volume in the GM of the aPFC had the smallest lesion in the group. This is in line with previous findings suggesting the recruitment of brain regions within domain-general networks is more likely for those with widespread damage in the language network (Kiran, Meier & Johnson, 2019).

Interpreting the function of the aPFC in relation to the iTalkBetter therapy is a more complicated matter considering the many roles this region plays in cognition. Neural adaption here may be related to the inhibition of speech responses, as hypothesised by Fridriksson and colleagues. Within the iTalkBetter therapy, participants practised naming a variety of words which included abstract items and those which were semantically related. This perhaps resulted in difficulties remembering the correspondence between pictures and target words and led to the need to inhibit responses which were incorrect. Although structural change in the aPFC did not correlate with treatment response, at the group level there was a significant decrease of all error types on the main outcome measure from pre- to post-therapy (see Chapter 1). Therefore, changes in this region may reflect an enhanced ability to inhibit the production of errors.

Another interpretation relates to the findings associating the aPFC with awareness of competence (Fleming, Weil, Nagy, Dolan & Rees, 2010). In iTalkBetter, feedback on whether the correct word was produced was provided for each trial and this may have improved self-monitoring skills, leading to better self-judgement of naming abilities. The aPFC has also been implicated in attention, specifically, controlling attention between inner thoughts and the external world (Burgess & Wu, 2013). It is possible that increased GM in this brain region reflects the participants' need to direct the mind to the repetitive and intensive external stimuli, which was the iTalkBetter therapy, rather than allowing the brain to wander and focus on inner thoughts.

5.2.3.4 The role of the right hemisphere

In the RH, increases in GM and WM were found in the PT (BA22), the RH homologue to Wernicke's area. In the dual stream account of language, proposed by Hickok and Poeppel (2007), in the ventral stream, phoneme discrimination, the first stage of auditory analysis, occurs bilaterally in the STG (PAC and PT) and the STS. The bilateral activations of these areas during speech processing has been supported by numerous studies with healthy individuals (Saur et al, 2010; Okada et al, 2010; Vigneau et al, 2011; Turken & Dronkers, 2011). On the other hand, the posterior part of the PT (Sylvian parietal-temporal junction, Spt) in the LH is believed to be the sensorimotor interface of the dorsal stream of language processing, connecting auditory and motor regions and projecting to the inferior frontal gyrus (IFG), the insula and the premotor cortex to enable speech articulation (Buchsbaum, Hickok & Humphries, 2001; Buchsbaum, Olsen, Koch & Berman, 2005; Hickok & Poeppel, 2015).

In accounts of spontaneous recovery following aphasic stroke, changes in the activation and in the structure of the right temporal lobe, around the region identified in this study, have been found in relation to both speech comprehension and speech production (Leff et al, 2002; Xing et al, 2016). In the study by Leff and colleagues, positron emission tomography was employed to identify if a laterality shift could be observed in 15 PWA who had damage to the pSTS, but had recovered their auditory processing of single words. The authors found that, in comparison to the eight control subjects, the aphasic individuals showed the same level of activity in pSTS but rather than the activity being located in the LH, it was in the RH. Furthermore, Xing and colleagues utilised lesion symptom mapping to compare GM volume between 32 PWA and two groups of matched controls to see if differences in structure related to neuroplasticity. Overall, the PWA had increased GM in the right temporo-parietal cortex in comparison to both control groups (healthy participants and those who had a stroke but had no history of aphasia). Additionally, increased GM volume for the aphasic group positively correlated with language outcomes, with clusters which included BA22 being associated with better repetition, naming, spontaneous speech and verbal working memory.

In the longitudinal structural imaging studies previously discussed, evidence of changes in RH temporo-parietal regions were also found. In Fleming and colleagues' study, increases in the GM of the right PT correlated with improved single word comprehension of trained items in a group of 25 PWA. Schlaug, Marchina and Norton found increased volume and number of fibres in the arcuate fasciculus (a major WM tract linking the IFG, the STG and primary motor cortex) of the RH in six participants with aphasia following Melodic Intonation Therapy (MIT). Wan and colleagues also used MIT in a group of 11 PWA and identified increases in the WM

of the RH in the IFG and PT which correlated with better speech fluency post-therapy.

The findings of increased GM and WM in the RH in this study provides further evidence for the recruitment of the RH in post-stroke aphasia. Whether this neuroplasticity was adaptive or maladaptive in relation to recovery is unclear based on the present data as changes were not associated with improvements. It is perhaps probable that the RH plays a differential role across individuals; contributing to recovery for some but hindering improvements for others, especially given the variability in hemispheric contributions to language processing in healthy populations (Catani et al, 2007). Similar to the findings in the LH, however, the majority of the participants showed neuroplasticity in both the GM (16 out of 17) and the WM (15 out of 17) of the RH. One of the participants (P3) showed decreases in volume across all three significant clusters identified in the analysis (in the GM and WM of the RH and the GM of the LH), suggesting general atrophy over the six weeks of therapy. The other participant (P19) showed decreases in volume in the WM of the RH only. As the size of P19's lesion was within one standard deviation of the mean of the group (therefore, was not abnormally small or large in comparison to the group), this provides tenuous evidence that lesion size did not affect RH structural changes.

In interpreting the reason for the neural adaptation observed in the RH following iTalkBetter therapy, a couple of explanations may be possible. Over the course of therapy, participants completed, on average, 17,538 individual lexical challenges (see Chapter 1) in which they had to name pictures. In the majority of these challenges, auditory cues were provided (whole word or phonemic cue). Therefore, the first interpretation relates to the role of the PT in speech comprehension. As the

initial stages of auditory processing occurs bilaterally, structural changes in the RH may reflect the recruitment of regions typically involved in language processing as a result of listening to the large number of auditory cues provided within iTalkBetter. Conversely, the second interpretation links to the role of the PT within the dorsal stream, connecting auditory and motor regions for speech articulation. The therapy required participants to repeatedly map the sound they heard to articulation, perhaps leading to having to rely on, and reorganise, RH homologues due to their extensive lesions in the LH (see Figure 5.13) (Friederici, 2012; Hickok & Poeppel, 2007; Wise, 2003).

5.2.3.5 Pre-therapy and post-therapy structural changes

In the post-hoc exploratory analysis, findings suggest there was natural atrophy in the right temporal lobe (in GM and WM) at the group level without therapy and the six weeks of iTalkBetter may have curtailed this general loss of volume. On the other hand, in the left aPFC the volume of GM and WM was stable over the pre-therapy block and iTalkBetter therapy seems to have induced hypertrophy in this brain region.

The natural atrophy seen in the right temporal lobe over the pre-therapy (or standard care) block reflects the neurobiological sequela of stroke which persists into the chronic stage, leading to further changes in structure and function over time. Although all of the participants in this study had LH lesions, RH homologues (for example, the PT) are structurally and functionally connected to the LH language network (Turken & Dronkers, 2011), and atrophy in these regions remote from the region of damage is caused by secondary processes such as diaschisis and

Wallerian degeneration (Brodtmann et al, 2020; Zhang, Zhang, Xing, Liang & Zeng, 2012).

Indeed, the Cognition and Neocortical Volume After Stroke (CANVAS) study specifically examined changes in total brain volume over a three-year period in 93 participants who had suffered from an ischaemic stroke (Brodtmann et al, 2021). The researchers found that post-stroke participants had greater total brain loss in comparison to a group of control subjects. Furthermore, those who additionally had cognitive impairments at three months post-stroke had greater atrophy than those with no cognitive impairments. The current findings therefore offer an important insight into possible avenues in which post-stroke secondary neurodegeneration could be curtailed using behavioural interventions. As increased brain atrophy and cognitive decline further impacts recovery from stroke, research which combines behavioural outcomes with neuroimaging methods is particularly important in identifying treatments which may reduce the vascular brain burden.

In the left frontal lobe, however, the hypertrophy observed over the therapy block is comparable to increases in volume in GM evident in healthy individuals when learning new skills (Draganski et al, 2004; Bezzola, Merillat, Gaser & Jancke, 2011; Mårtensson et al, 2012). This suggests there was no general atrophy in this region without therapy but completing the iTalkBetter task promoted neuroplasticity. This is in contrast to the CANVAS study in which the frontal lobes were found to be especially vulnerable to greater brain atrophy post-stroke. This difference in results could be due to a number of factors, such as sample size and heterogeneity in lesion location.

5.2.3.6 Limitations

The findings in the iTalkBetter study identified changes in GM and WM over a six-week intervention period, over and above changes during a six-week control period. These structural changes, however, were not related to treatment outcomes but were an overall effect of completing the therapy. The reason the neuroplasticity observed was not associated with therapy response may be due to the small sample size (n=17) included in the analysis. Although the original plan for this clinical trial had been to include as many participants as possible in the MRI component of the study, the COVID-19 pandemic severely affected the ability to complete scanning for many of the PWA. Adding in the additional covariates of behavioural change and dose in the statistical models reduces the degrees of freedom and, therefore, due to the sample size, analyses 2 and 3 may have been underpowered. Furthermore, the small sample size makes it difficult to generalise the results across the aphasic population due to the likelihood of increased false negative and false positive results when employing a mass-univariate approach, as utilised in this study (Lorca-Puls et al, 2018).

Another limitation of all VBM methods is interpreting exactly what the changes in volume demonstrate at a cellular level. It is possible experience-dependent neuroplasticity in GM and WM reflects changes in neuronal size, axonal or dendritic arborisation, or neuropil (Mechelli, Price, Friston, & Ashburner, 2005). It is not clear in the current study how the increases in volume relate to the microstructure of the brain. Regardless of the underlying neural mechanisms, however, VBM is a powerful, unbiased, technique for investigating volumetric differences across the whole brain and, as shown in present results, it is able to identify small-scale differences in GM and WM across individuals over time.

5.2.3.7 Summary and future considerations

The current study adds to the literature by providing evidence of structural plasticity in the domain-general network and in right hemisphere language homologues following anomia treatment. The within-subject design of the study overcomes the frequently reported methodological limitation of a lack of control group (Thomas & Baker, 2013), as participants acted as their own controls, reducing between-subject variance and regressing out structural changes related to time (for example, due to natural degeneration). The study also utilised a methodology previously used in the aphasia research and identified plausible regions of change, supporting the use of longitudinal VBM as an analysis tool to identify structural neuroplasticity in PWA.

Due to the lack of findings of change related to therapy response, however, future work that examines neural adaptation at an individual level may offer a more comprehensive view on neuroplasticity mechanisms and how they relate to recovery. This could be achieved by employing a multivariate analysis approach and would be especially insightful considering the research indicating differential patterns in recovery, possibly related to factors such as lesion size and location (Wilson & Hula, 2019; Stockert, 2020; Fridriksson et al, 2006; Vitali et al, 2007). Additionally, research both in healthy subjects and in aphasic individuals suggests changes in structure and function may differ over time, with immediate treatment effects not always mirroring long-term neuroplasticity (Draganski et al, 2004; Menke et al, 2009). For example, in the study by Draganski and colleagues, neural adaption was investigated following three months of juggling practice in a group of healthy participants who were inexperienced jugglers. Although immediately following this training period increases in GM were observed, three months after the end of practice, the increase in GM had diminished in all subjects. In future work examining

structural neuroplasticity following aphasia therapy, a similar follow-up scanning session would be useful in elucidating whether treatment effects are maintained or if brain tissue renormalises following the cessation of intervention.

5.3 Chapter 3: Investigating changes in task-related brain activity in response to iTalkBetter therapy.

Aim: To investigate whether iTalkBetter therapy induces changes in the language and/ or cognitive networks, identified using fMRI scanning during speech production tasks, in persons with chronic aphasia.

Hypotheses:

(1) iTalkBetter therapy will induce changes in functional activation in the language and/ or cognitive networks, identified using speech production tasks (bilateral frontal and/ or temporal lobes).

(2) Changes in task-related activation will correlate with percentage change on trained items on the WRT following therapy.

(3) Dose of therapy (number of hours completed) will correlate with changes in task-related activation.

5.3.1 Experimental procedures

5.3.1.1 Participants

In the following analyses, data from the same subset of 17 participants from Chapter 2 are presented.

5.3.1.2 Design

A repeated measures design was used, with three scanning time points: two pre-therapy (T2 and T3) and one post-therapy (T4). A description of the scanning protocol is detailed in the Methods.

5.3.1.3 Data pre-processing

The data was smoothed using an isotropic kernel of 8mm full-width half-maximum and the prior pre-processing steps are discussed in the Methods.

5.3.1.4 Statistical analyses

17 contrast images from the single subject first-level analysis (see Methods) were entered into three separate whole brain, second-level, one sample t-tests in SPM12. Each contrast image represented change in activation at post-therapy (T4) in comparison to an average of pre-therapy (T2 and T3) on the three speech production tasks.

Analysis 1: To investigate change in activation as an overall effect of partaking in the therapy program, the 17 contrast images were entered into the second-level analysis with no covariates.

Analysis 2: To identify if there were changes in activation which correlated with response to therapy, percentage change on trained items on the WRT from pre-therapy (T3) to post-therapy (T4) was entered as the dependent variable in the model (Figure 5.18a).

Analysis 3: Dose (number of hours of therapy completed) was entered as the dependent variable in the third model to ascertain if there changes in activation which correlated with the amount of therapy participants achieved (Figure 5.18b).

As in Chapter 2, exposure was entered as a covariate of no interest in both analyses 2 and 3 due to differences in exposure to trained items (unrelated to dose) across participants (see Methods and Chapter 1). However, as with the structural pre-processing, the within-subject design of the first-level analysis accounted for

baseline behavioural and demographic between-subject effects, so no further regressors were entered into the models. The statistical voxel-level threshold was set at p<0.001 uncorrected, and to correct for multiple comparisons, significant clusters are reported at the Family Wise Error (FWE) p<0.05 threshold.

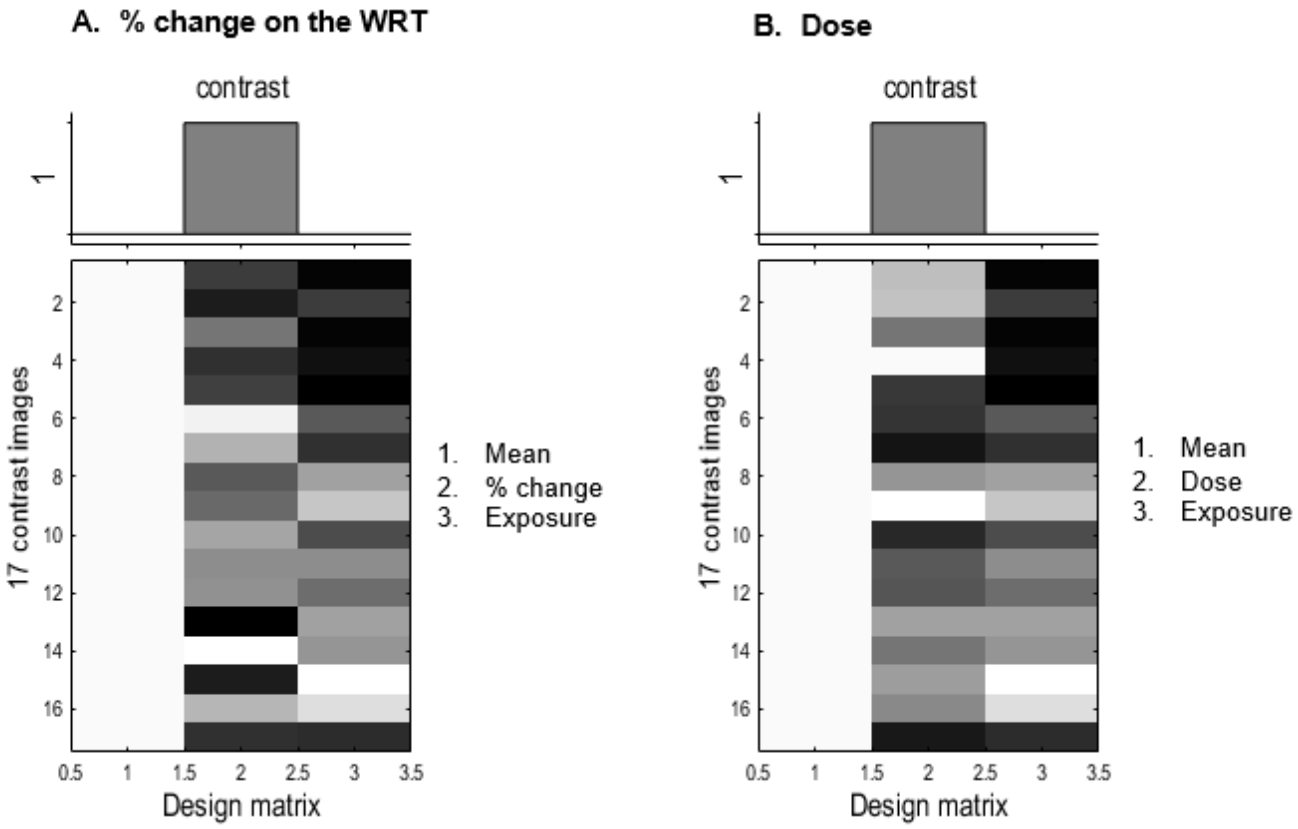


Figure 5.18 Design matrices used in the simple linear regression models for Analyses 2 and 3. (A) Change on trained items on the WRT (%); (B) Dose (number of hours of therapy completed). Both models also included exposure as a covariate of no interest.

5.3.2 Results

5.3.2.1 Behavioural results

Figure 5.19 below displays performance on the fMRI tasks at each time point. A repeated measures ANOVA (one factor: time; three levels: T2, T3 and T4), found no significant effect of time on the number of overall speech responses given for the three tasks, F(2,15)=1.5, p=0.24 (T2: M=48, SD=12; T3: M=48, SD=12; T4: M=49, SD=10). To investigate if there was a change in performance on each of these tasks over time, three repeated measures ANOVAs were completed which all had one factor (time) with three levels (T2, T3 and T4). There was no significant change over time in accuracy on either the verb naming (VN) or sentence naming (SN) tasks (VN: F(2,15)=3.12, p=0.06; SN: F(2,15)=1.12, p=0.34). There was, however, a significant effect of time for the object naming task (ON), F(2, 15)=7.01, p=0.003, which was driven by an improvement from T3 (M=8, SD=5) to T4 (M=10, SD=7) at the group level, *t*(16)=2.59, p=0.02. There was no significant change in performance on the ON task at T2 (M=8, SD=5) in comparison to T3, *t*(16)=-0.7, p=0.5.

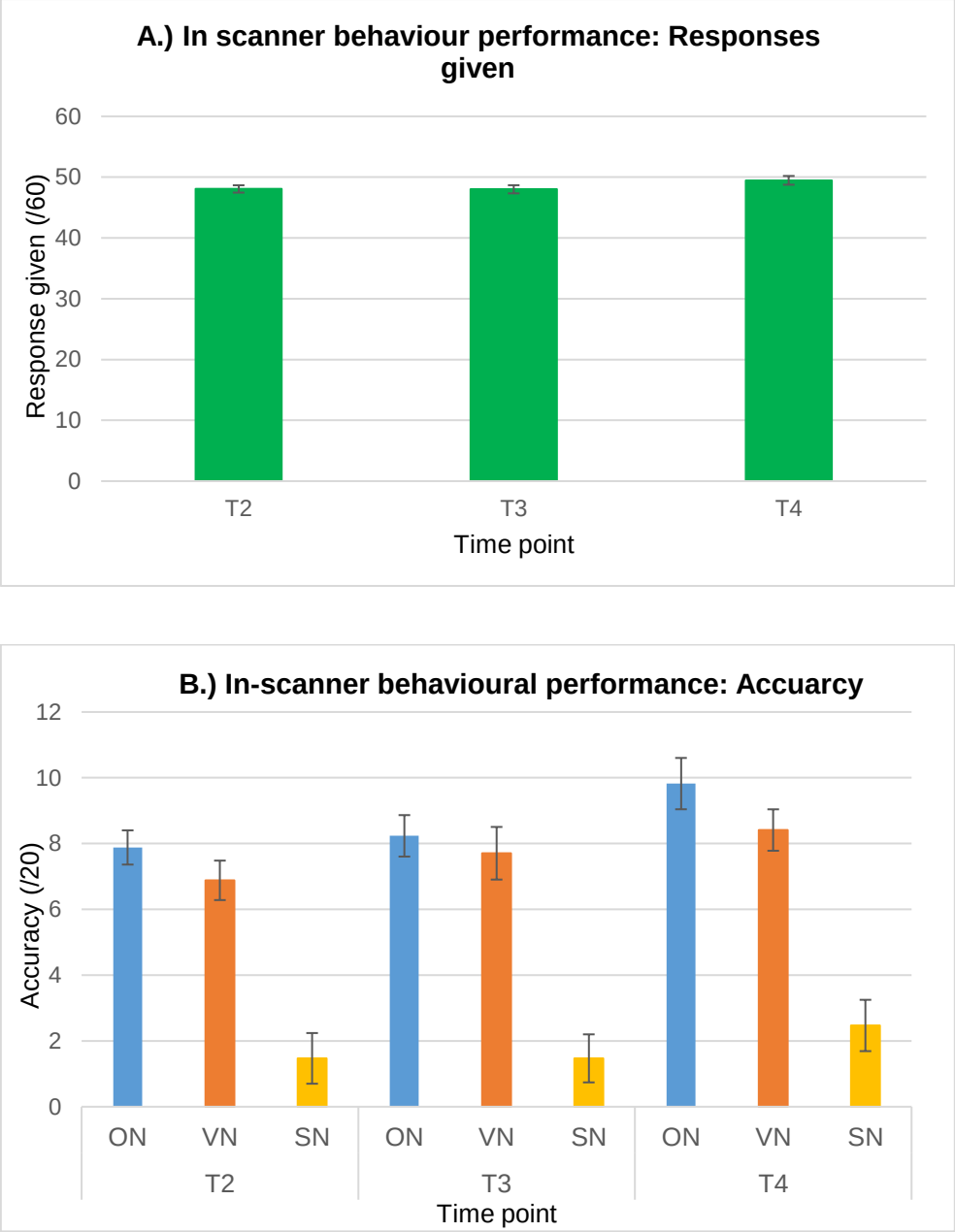


Figure 5.19 In-scanner behavioural performance at T2, T3 and T4. (A) Responses given for the three speech tasks; (B) Performance accuracy on the three speech tasks (ON = object naming; VN = verb naming; SN = sentence naming). Error bars are within-subject standard error of the mean.

In the first level design matrix, responses from all three speech tasks were analysed together at pre-therapy (T2 and T3) and post-therapy (T4). However, as the participants who took part in this study varied in the severity of their aphasia, some participants were unable to produce correct responses for all tasks at all of the time

points. To enable the inclusion of these participants within the fMRI analysis, correct and incorrect responses were collapsed together and coded as 'speech responses' (as described in the Methods).

5.3.2.2 Analyses 1 and 2

Whole brain fMRI analysis investigating change in functional activity at post-therapy (T4) in comparison to pre-therapy (T2 and T3) with no added covariates of interest did not reveal any significant activations. Additionally, adding change in trained items on the WRT as a dependent variable in the model (with exposure as a covariate of no interest), did not unveil any significant changes in activation at post-therapy. These results do not support Hypotheses 1 and 2, suggesting therapy alone and response to therapy were not correlated with changes in task-related functional activation patterns in this group of participants.

5.3.2.3 Analysis 3

In the third analysis, dose of therapy was included in the model as the dependent variable, with exposure again as a covariate of no interest. This whole brain analysis revealed two significant clusters which increased in task-related activation from pre-therapy to post-therapy in both the left (n=1) and the right hemispheres (n=1), supporting Hypothesis 3, and are displayed in Figure 5.20 and Table 5.5. In the LH, the first two peak voxels in the cluster were in the insula (BA13), with a further peak in the superior temporal sulcus, but which visually corresponds with the primary auditory cortex (PAC, BA41) due to the larger anatomical area of this brain region in the LH (Penhune, Zatorre, MacDonald & Evans, 1996). In the RH, the peak voxel was in the PAC (BA41), with further peaks extending into the insula (BA13).

In Figure 5.21, the scatterplots illustrate activation changes at post-therapy in relation to dose for eignevalues from 5mm spheres around the peak voxel in the clusters for each participant. The third scatterplot displays the correlation between activation in the LH and activation in the RH, illustrating differential patterns of changes between participants. Eight participants had increased activity bilaterally following therapy (P2, P3, P4, P5, P6, P16, P19, P24); three participants had an increase in activity in the right hemisphere but a decrease in activity in the left hemisphere (P18, P23, P27); oppositely two participants showed increased activation in the left hemisphere but decreased activation in the right hemisphere (P21, P22); and finally, four participants had decreases in activity in both hemispheres following therapy (P17, P20, P25, P26). Additionally, as the cluster in the RH overlapped with the area of change identified in the structural imaging, correlational analysis was completed to ascertain whether task-related activity changes were associated with volume changes. This revealed no relationship between increases in post-therapy task-related neural activation and post-therapy increases grey matter or white matter (GM: $r(17)=0.12$, $p=0.66$; WM: $r(17)=-0.12$, $p=0.65$) (Figure 5.22).

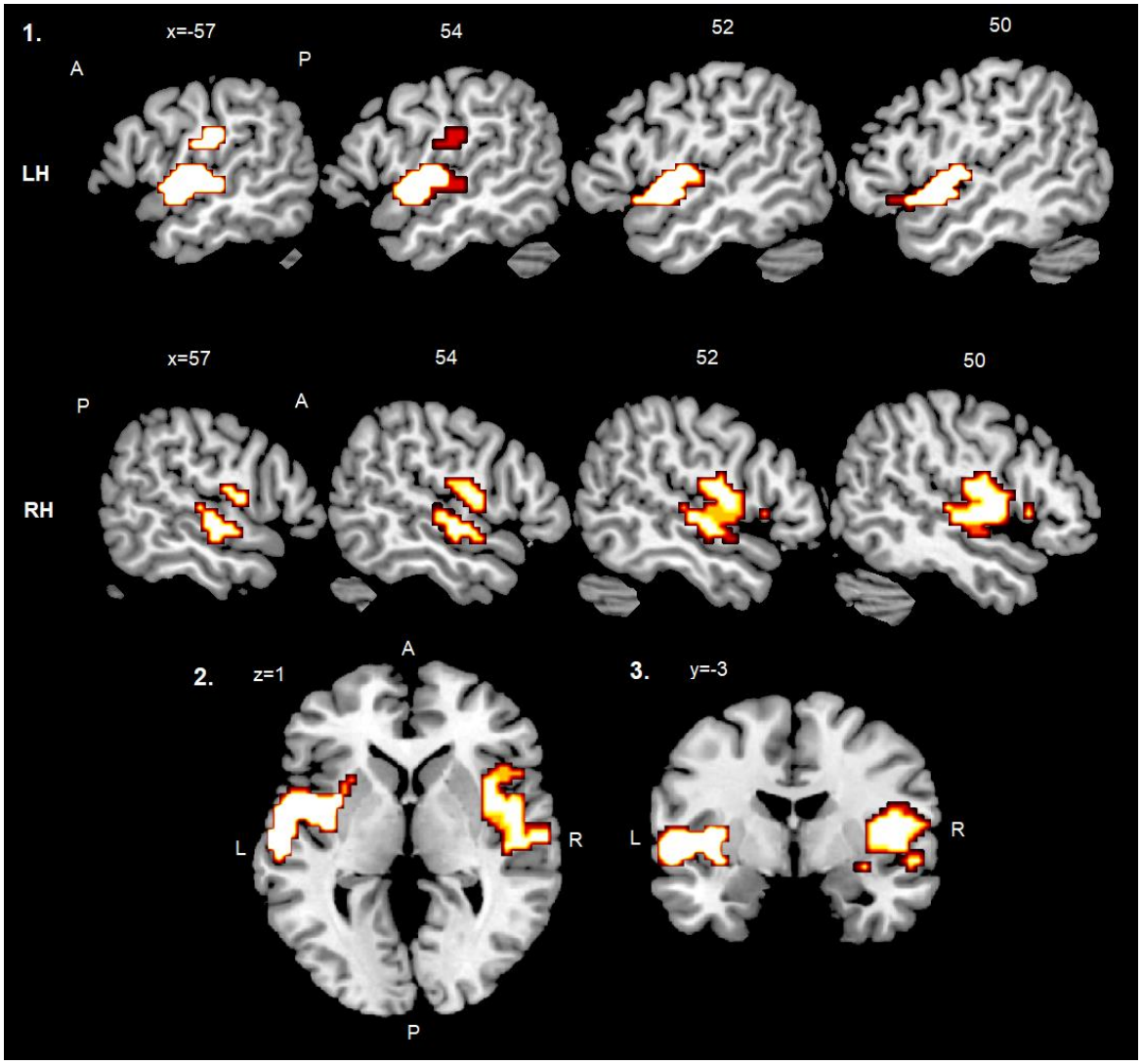


Figure 5.20 Sagittal (1), axial (2) and coronal (3) views of higher relative activation at post-therapy. LH = left hemisphere; RH = right hemisphere. A = anterior; P = posterior; L= left; R = right. Numbers represent MNI coordinates.

Left hemisphere	***k***	***T***	***x***	***y***	***z***
Insula	339	5.72	-30	14	-7
		5.37	-48	5	-4
Primary auditory cortex (PAC)		4.97	-57	-1	2
Right hemisphere	***k***	***T***	***x***	***y***	***z***
Primary auditory cortex (PAC)	378	9.40	60	-13	-1
Insula		5.73	48	-4	2
		5.63	36	5	-1

Table 5.5 Significant clusters with increases in activation at post-therapy. The numbers represent MNI coordinates of the first three peak voxels in the clusters; 'k' is cluster size; and 'T' is the t-value.

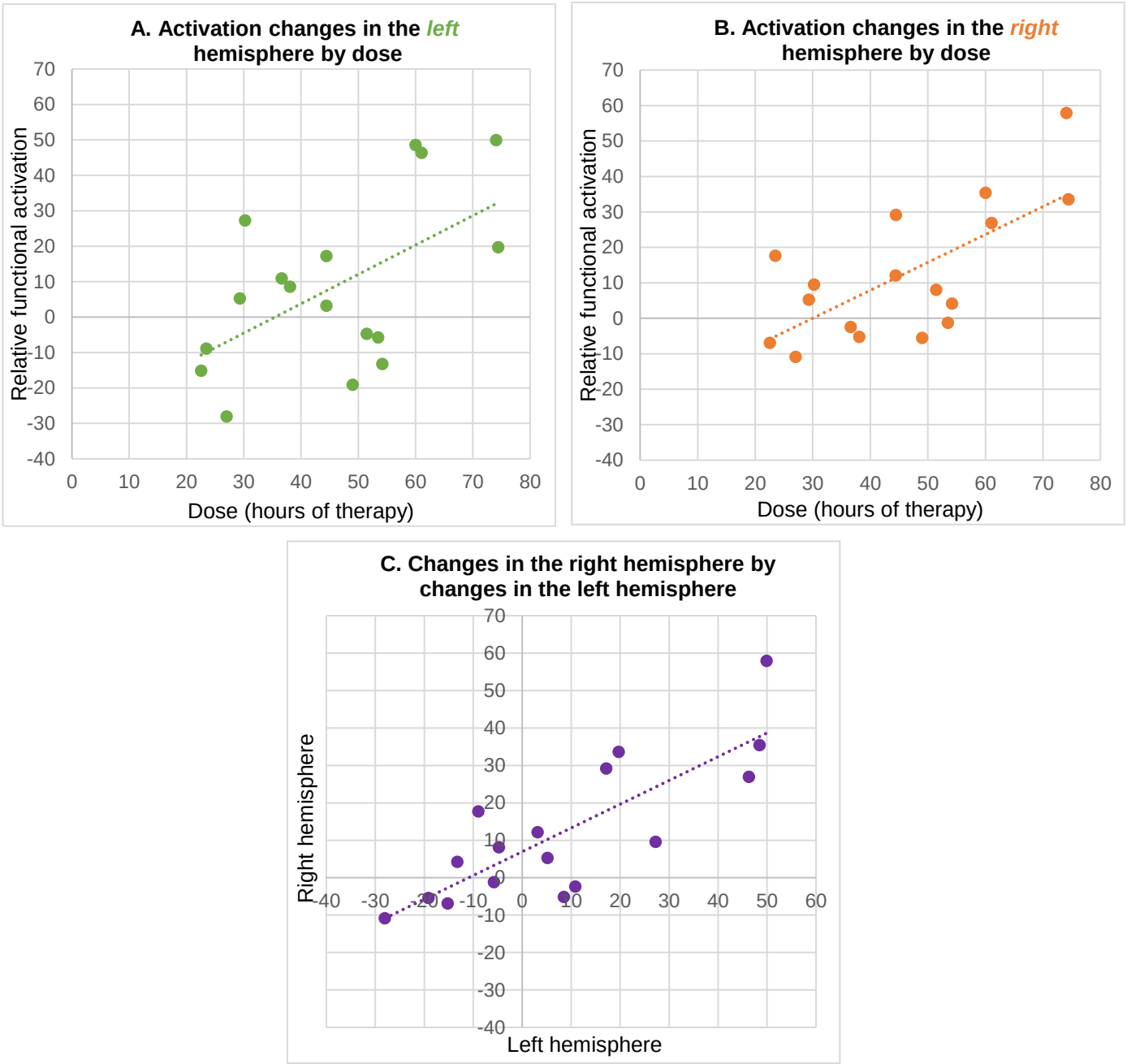
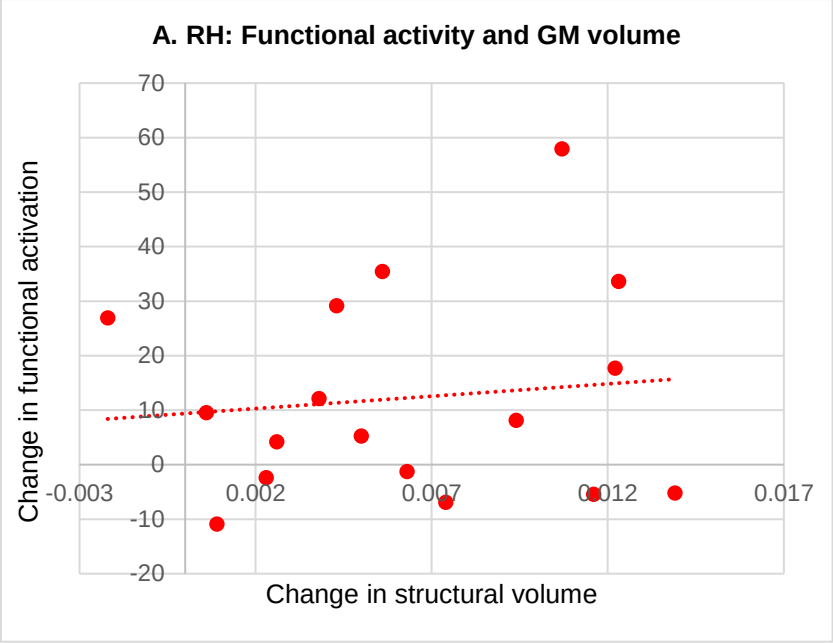


Figure 5.21 Scatterplots showing changes in activity which correlated with dose at post-therapy. (A) Activation changes in the left hemisphere; (B) Activation changes in the right hemisphere; (C) Changes in the left hemisphere by changes in the right hemisphere. Relative activation values are eigenvalues from 5mm spheres around the peak voxel in each of the clusters.



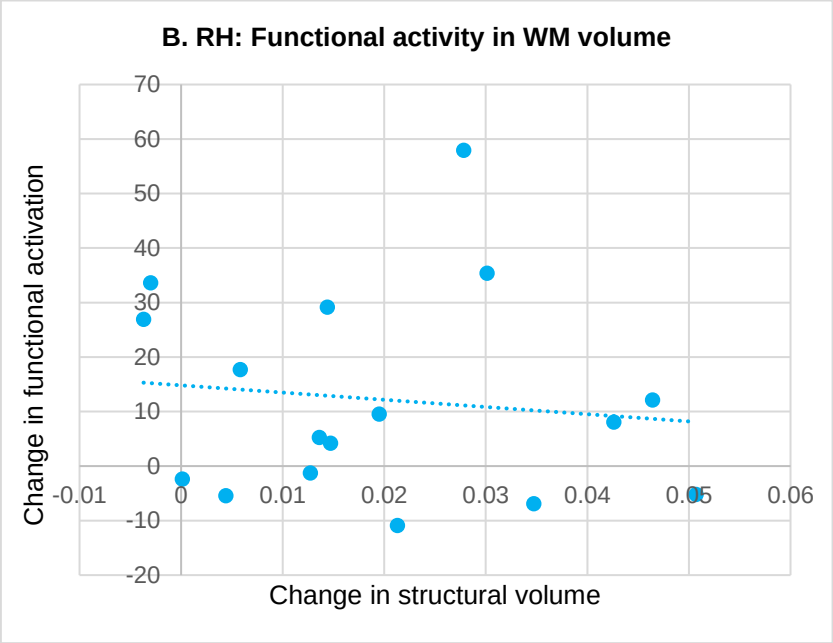


Figure 5.22 Scatterplots showing changes in functional activity and changes in structural volume in the right hemisphere. (A) Changes in functional activation and changes in grey matter; (B) Changes in functional activation and changes in white matter. Relative activation values are eigenvalues from 5mm spheres around the peak voxel in each of the clusters. Volume values are also eigenvalues from 5mm spheres around the peak voxel in each cluster.

5.3.3 Discussion

5.3.3.1 Summary of results

This chapter aimed to identify whether there were any changes in task-related neural activity in the language and cognitive networks in a group of participants with chronic aphasia following six weeks of iTalkBetter therapy. Whole brain fMRI analyses did not support Hypotheses 1 and 2, as no significant areas of activation were identified in relation to simply taking part in therapy or in relation to improvement following therapy. In analysis 3, however, significant increases in task-related activity were found to correlate with dose in both the left and the right hemispheres. This supports Hypothesis 3 and suggests that a higher dose of therapy led to increases in activation bilaterally in areas supporting both auditory analysis (primary auditory cortex, BA41) and speech production (insula, BA13).

Interestingly, in the RH, the area of activation change overlapped with the cluster identified in the structural imaging analysis. Although in the structural imaging this was an overall effect of therapy and was not related to dose of therapy. Due to this, predictably, these changes in grey matter and white matter volume in the RH were not correlated with changes in task-related activation.

5.3.3.2 Comparison with previous findings

In imaging studies with healthy participants, longitudinal changes in neural activity have been observed following training in tasks such as auditory discrimination and musical practice (Tremblay & Kraus, 2002; Herdener et al, 2010). In language acquisition, functional changes have also been found post-training in artificial and foreign language learning (Xue, Chen, Jin & Dong, 2006; Barbeau et al, 2017). For example, Xue and colleagues assessed changes in task-related activity following a

six-week, logographic, artificial language training programme with 12 individuals. Training consisted of two weeks of practising the visual form of the novel characters, followed by two weeks of phonological training and two weeks of semantic training. After phonological and semantic training, increases in activity in areas of the language network were identified, including the left inferior frontal gyrus and the left insula. Additionally, Barbeau et al completed a 12-week French language learning programme with 14 participants and, following training, found increased activation in the left inferior parietal lobule (including the supramarginal gyrus) when reading French sentences, but not when reading English sentences.

Studies investigating task-related activity changes following aphasia therapy have identified mixed results, with both increases and decreases in the LH and RH being observed (Kiran & Thompson, 2019; Li, Mukadam & Kiran, 2022). This diversity in findings is perhaps revealing of the heterogeneity of PWA across intervention studies, differences in therapy approach and assessment methodologies, as well as the large and complex networks involved in language processing. Indeed, many studies have identified differential bilateral patterns of neural activity changes following treatment programmes for anomia, with both increases and decreases bilaterally correlating with therapy response (van Hees, McMahon, Angwin, de Zubicaray & Copland, 2014; Nardo, Holland, Leff, Price & Crinion, 2017; Fridriksson, Morrow-Odom, Moser, Fridriksson & Baylis, 2006; Richter, Miltner & Straube, 2008; Abel, Weiller, Huber, Willmes & Specht, 2015).

The present results add to those above by providing evidence of increased bilateral task-related activity in regions of the language network following word retrieval practice. At the group level, the bilateral activation observed was an increase in activity from pre- to post-therapy, however, there was variability between

participants, with some showing increased activity in the LH or RH only, and some showing bilateral decreases. This reflects the broad range of neural adaption patterns observed in aphasic individuals and, although not seeming related to lesion size (see Table 4.1 in Methods), may be due to a range of factors such as lesion location, severity and level of impairment, and premorbid differences in the language network (Kiran & Thompson, 2019). Also, as with the structural imaging, post-therapy neuroplasticity did not correlate with treatment outcomes, indicating between-subject differences in neural adaption in relation to recovery and the adoption of individual strategies to improve word retrieval skills. In this fMRI analysis, however, task-related activation was associated with the amount of therapy participants completed, suggesting higher doses lead to increased activity bilaterally. Although Marcotte and colleagues (2018), examined changes in activation following differential intensity schedules of therapy, no studies, to my knowledge, have investigated the relationship between treatment dose and changes in neural activity in the aphasia literature. In structural imaging, however, there is evidence in healthy individuals that the amount of training in tasks such as phonetic transcription results in differences in brain volume in regions of the language network (Golestani, Price & Scott, 2011).

5.3.3.3 The Insula

In the LH, the first two peak voxels in the cluster which increased in activation over the therapy block were in the insula (BA13); as was the second peak voxel in the cluster in the RH. The insula is connected to the planum temporale and the premotor cortex and is thought to mediate speech articulation in the dorsal stream of language processing (Buchsbaum, Hickok & Humphries, 2001; Buchsbaum, Olsen, Koch & Berman, 2005; Hickok & Poeppel, 2015). The bilateral insulae, however, have also

been implicated in receptive language tasks, in addition to expressive tasks, in healthy individuals, suggesting a role in the co-ordination of higher cognitive aspects of the language network (Skipper-Kallal, Lace, Xing & Turkeltaub, 2017; Oh, Duerden & Pang, 2014). The anterior insula is also part of the salience (cingulo-opercular) network which is believed to be involved in cognitive processes such as behavioural inhibition and attentional control (Dosenbach et al, 2008; Stockert et al, 2020; Geranmayeh, Brownsett & Wise, 2014; Seeley, 2019; Burgess & Wu, 2013).

In fMRI studies examining spontaneous recovery following aphasic stroke, task-related changes in activity have been observed in both the right and the left insulae (Szaflarski et al 2011; Allendorfer, Kissela, Holland & Szaflarski, 2012; Stockert et al, 2020). For example, in the study by Allendorfer and colleagues, patterns of language activation during an overt verb generation task were investigated using fMRI paradigms of both an event-related design (to isolate activations relating to semantic processing, articulation and auditory processing) and a block-design. In the event-related fMRI paradigm, auditory processing and articulation were associated with increased activity in the right insula in comparison to a group of 32 healthy controls. Furthermore, in Stockert and colleagues' study, increased activity in the left insula (but not the right insula) correlated with greater improvements in comprehension and production skills for those with temporo-parietal damage.

Studies investigating therapy-induced functional neuroplasticity have also found varied results in terms of the contribution of the insula, although these findings have predominantly been in the RH. In the study by Nardo, Holland, Leff, Price and Crinion (2017), increased activation in the right anterior insula correlated with immediate and long-term improvements in naming following phonological therapy. Additionally, in a study using positron emission tomography, Raboyeau and

colleagues (2008) found that increased cerebral blood flow in right anterior regions and in the right insula correlated with improved lexical retrieval following word finding therapy in a group of 10 PWA. On the other hand, following two weeks of constraint-induced aphasia therapy with 16 PWA, Richter, Miltner and Straub (2008) identified decreased activity in the right insula which was associated with higher behavioural improvement scores (spontaneous speech, auditory comprehension and sentence comprehension) in an fMRI task employing word stem completion.

The current study provides further evidence of functional change in the insula following word finding therapy in aphasic individuals, and indicates a bilateral recruitment of this brain region at the group level related to the amount of practice rather than treatment outcomes. The recruitment of the insulae following iTalkBetter therapy may be due to the role it plays in speech articulation, with repeated word finding practice over the six weeks leading to increased activation. Alternatively, changes may have occurred in this region because of its connection to the cingulo-opercular network, indicating improved attention, awareness of competence, or inhibition of incorrect responses, as discussed in Chapter 2. The final interpretation relates to the bilateral activation of the insulae during speech perception tasks in healthy individuals and, as the PAC was also identified as an area of change, increased neural activity in the insulae could be related to changes in auditory analysis skills.

5.3.3.4 The primary auditory cortex

In the RH, the peak voxel in the cluster which increased in activation over the therapy block was in the PAC (BA41); as was the third peak voxel in the cluster in the LH. The PAC is located on the superior temporal gyrus (STG) and, as mentioned

in the discussion of Chapter 2 and in the introduction, the early stages of auditory analysis occurs bilaterally in healthy individuals in the STG and superior temporal sulcus (STS). (Saur et al, 2010; Okada et al, 2010; Vigneau et al, 2011; Turken & Dronkers, 2011). For example, in Vigneau and colleagues' meta-analysis of 59 studies, bilateral activations in the PAC and the posterior part of the STS (pSTS) were observed in healthy participants when listening to speech sounds as opposed to non-speech sounds, indicating interhemispheric cooperation.

In accounts of spontaneous recovery following aphasic stroke, changes in the activation in the right temporal lobe, in and around the region identified in this analysis, have been identified. For example, in an fMRI study by Skipper-Kallal and colleagues (2017), increased activation in the right STG and STS, including the PAC, was observed during an overt naming task in 39 PWA, in comparison to a group of healthy control subjects. Furthermore, Fridriksson, Baker and Moser (2009) completed an fMRI study with 11 individuals with chronic aphasia and found that correct responses of concrete nouns were associated with increased activations in the right hemisphere. The areas implicated included the planum temporale (PT, BA22) and the supramarginal gyrus (BA40) which are two regions adjacent to BA41. Szaflarski and colleagues (2011) also found increased activation in the left STG in four PWA, in comparison to a group of healthy control participants, during a written picture matching fMRI paradigm.

Research investigating therapy-induced activation using fMRI have also identified changes in the STG in response to word finding therapy, though again differential findings have been observed. In a study by Marcotte and Ansaldo (2010) with two participants, one with primary progressive aphasia and the other with post-stroke aphasia, Semantic Feature Analysis (SFA) therapy was found to increase activation

in the STG in the LH when retrieving trained verbs for the PWA. However, in the more recent study by Marcotte and colleagues (2018) which employed Phonological Component Analysis (PCA) to compare different intensity schedules of therapy, the non-intensive (but not the intensive) programme resulted in increased activation bilaterally in the STG. Alternatively, Abel, Weiller, Huber, Willmes and Specht (2015) completed a word finding therapy programme using semantic and phonological cueing hierarchies with 14 PWA. Following therapy, although participants displayed differential patterns of neuroplasticity, those with both phonological and semantic impairments showed decreases in activity in the left STG.

The results from this task-based fMRI analysis provides evidence of bilateral recruitment of regions in the STG following iTalkBetter therapy which increased with the amount of therapy completed. In the current study, this activation was in the PAC so it is likely that changes here relate to the auditory cueing paradigm employed in the therapy programme. As mentioned in Chapters 1 and 2, participants completed, on average, over 17,000 individual trials, the majority of which would have had an auditory cue. It is, therefore, not surprising that neuroplasticity was observed in regions underlying auditory analysis. Additionally, as the therapy provided feedback on whether the correct response was given, changes in the PAC may be connected to participants' improving their self-monitoring skills by listening to their own speech responses.

5.3.3.5 Limitations

The findings in the iTalkBetter study identified changes in task-related activation at post-therapy in comparison to two pre-therapy time points, but a couple of limitations must be addressed. Although the sample is this study is larger than many of those

reported above, it was still relatively small (n=17) due to the restrictions surrounding the COVID-19 pandemic. Small sample sizes can lead to a lack of statistical power, reducing the replicability of results across the aphasic population, and may be why changes in task-related activity relating to treatment response were not identified. PWA are greatly diverse in terms of language and cognitive profiles, and lesion characteristics, and in this study the participants' size and location of damage in the LH was highly variable. Due to this, there may have been additional regions in the LH that did undergo neuroplasticity but were not identified in the analysis as they lacked power at the group level, resulting in possible Type 2 errors.

The tasks completed in the fMRI paradigm were also not specific to the treatment programme as they did not include words trained in the iTalkBetter therapy. They were, however, speech tasks and involved different levels of language processing (object, verb and sentence production) but these tasks were analysed as one. The participants in the study also found the tasks difficult and, due to this, correct and incorrect responses were analysed together. It has been suggested that there are different neural patterns which relate to correct responses versus speech errors and the statistical approach employed may have masked out these differences (Meinzer et al, 2013; Fridriksson et al, 2006).

5.3.3.6 Summary and future directions

The current study adds to the literature by providing evidence of bilateral changes which were related to dose of therapy in the insula and the PAC following six weeks of iTalkBetter word retrieval intervention. The post-therapy neuroplasticity was found in regions involved in speech perception and speech articulation, processes that are likely to be more homogenous across individuals in comparison to the word retrieval

stages of phonological and semantic processing. Also of note, the cluster identified in the LH was in an area which some participants had damage to following their stroke (see lesion overlap map in Chapter 2). As this temporal area did increase in neural activity, it suggests that although there is reduced grey and white matter density in this region, the remaining structural architecture is still functioning. Furthermore, the comparison of two pre-therapy baseline scans with a post-therapy scan gives assurance that these changes observed were due to treatment rather than between-scan variability.

Due to the between-subject differential findings in both the left and the right hemispheres and because no changes were found which related to treatment response, future work which employs a multivariate approach, taking into account intrinsic differences (such as lesion location, as discussed in Chapter 2), may give a more detailed insight into functional neuroplasticity at an individual level (Wilson & Hula, 2019; Kiran & Thompson, 2019). Additionally, as aphasia is a network disorder, a further consideration would be to examine changes in functional connectivity using connectivity-based analyses (Kiran & Thompson, 2019). This could be achieved either by using the task-based data reported in this thesis, or by using the resting-state fMRI imaging completed at each of the three scanning time points. Finally, due to the numerous studies reporting differential neural activation patterns which have found to predict response to specific treatment programmes, the task-based or the resting state data could be used to identify if pre-therapy functional activity is associated with response to iTalkBetter (Menke et al, 2009; Van Hees, McMahon, Angwin, de Zubicaray, & Copland, 2014; Fridriksson, Morrow-Odom, Moser, Fridriksson & Baylis, 2006; Vitali et al, 2007).

6 General discussion

6.1 Summary of key results

This thesis investigated the efficacy of the iTalkBetter therapy app in improving word retrieval skills in persons with chronic aphasia. Longitudinal structural and functional MRI analyses were also reported for a subset of participants to examine neuroplasticity in response to the therapy.

The three broad aims of this thesis were:

(1) **Chapter 1**: To test the effectiveness of iTalkBetter in a group of participants with chronic aphasia.

(2) **Chapter 2**: To investigate whether iTalkBetter induced regional changes in brain tissue using structural MRI.

(3) **Chapter 3**: To explore therapy-driven, task-related activity changes following iTalkBetter using functional MRI.

In **Chapter 1** I investigated the efficacy of iTalkBetter in a repeated measures trial with 27 participants with chronic post-stoke aphasia. All participants had anomia as part of their aphasia, as demonstrated by impaired single word naming. The results showed that iTalkBetter was effective at improving word retrieval as participants made large and significant gains in single word naming on the WRT (for items that were trained) which were maintained 12 weeks following the cessation of therapy. No comparable improvements were found for untrained items at the group level, suggesting an item-specific learning effect which is often reported for impairment-based interventions (Webster, Whitworth & Morris, 2015; Wisenburg & Mahoney, 2009; Raymer et al, 2008; Nickels, 2002).

Over the six-week intervention block, participants completed a large dose of therapy (45 hours), demonstrating that the therapy was engaging and well-tolerated. These findings support the use of computerised self-led therapies in delivering treatment in aphasia rehabilitation, and contribute to the evidence indicating a high dose of therapy is required to improve language functions (Brady et al, 2016; Zheng, Taylor & Lynch, 2016; Des Roches & Kiran, 2017; Palmer et al, 2019; Bhogal et al, 2003; Brady et al, 2022). This is a particularly important finding as currently the provision of speech and language therapy following a stroke is severely limited (Clarke et al, 2018; Palmer, Witts & Chater, 2018). Identifying cost-effective and efficacious methods of delivering high-dose rehabilitation are required to support those with aphasia.

Furthermore, significant gains were observed for both trained concrete and trained abstract items, suggesting the iTalkBetter therapy was effective in improving both word types. This result adds to the extremely limited literature base investigating the treatment of abstract words using traditional therapy techniques (Renvall & Nickels, 2019), and supports the use of these intervention approaches in improving the retrieval of more abstract items. As abstract words are integral to the ability to express emotions and complex needs (Worrall et al, 2011), are frequent within the English language (Renvall, Nickels and Davidson, 2013), and are often impaired in PWA (Walker & Hulme, 1999; Conroy, Snell, Sage & Lambon Ralph, 2012; Nickels & Howard, 1995; Kiran, Sandberg & Abbott, 2009), therapies which include words across the concreteness continuum are vital to ensure maximal effectiveness of aphasia therapy. In future work, it will be necessary to continue to assess the value of employing impairment-based interventions in targeting more abstract items.

Participants also significantly improved the number of trained items and the number of unique information carrying words they produced in the spoken picture description task. This demonstrates a generalisation of learning effects from the single word naming therapy to a less constrained task. As SPD is a similar task to single word picture naming, this across-level generalisation is an instance of near transfer (Subedi, 2004; Conroy, Sage and Lambon Ralph, 2009c). Previous studies investigating across-level generalisation following anomia intervention have found a lack of evidence indicating that improvements at the single word level generalise to connected speech (Nickels, 2002, Wisenburn & Mahoney 2009; Webster, Whitworth & Morris, 2015; Marshall et al, 2018). Improvements in the SPD task following iTalkBetter therapy are therefore a positive finding, however, it must be noted that no correlations were found which linked improvements on the single word naming test (WRT) to improvements in SPD. This is perhaps due to individual differences in the ability to generalise learning across communicative tasks and future investigations, using the data collected in this study, could examine whether there were any behavioural factors which promoted or impeded generalisation (Best et al, 2013; Holland, John & Woollams, 2018).

Across the behavioural findings discussed above, there was significant variability in treatment response at the individual level. Investigating the factors which may explain some of this variance is clinically important to enable evidence-based decisions regarding treatment. In iTalkBetter, for both the SPD and the WRT, a combination of baseline factors (both speech production and comprehension measures) accounted for a large amount of the variability, providing some indication of who benefitted most from the therapy. However, the analyses employed were purely explanatory and do not offer a predictive value for who is most likely to

respond to this specific therapy. Previous research which has used explanatory and predictive modelling have found that models have significantly less predictive power than explanatory power, possibly due to overfitting of the data (Aguilar et al, 2018). As PWA devote considerable time and effort to speech and language interventions, future analyses which investigate whether individual response to iTalkBetter (possibly also incorporating brain data) can be predicted would be clinically beneficial.

In **Chapter 2**, I explored structural neuroplasticity in response to iTalkBetter treatment in a subgroup of 17 participants who took part in the study. This analysis was completed by comparing structural changes over the six-week pre-therapy period, to the six-weeks of iTalkBetter intervention. Changes were found which were associated with taking part in the treatment programme, although these changes in brain tissue were not found to be related to improvements following intervention or the dose of therapy achieved.

Increases in volume were observed bilaterally in the grey matter of the left anterior prefrontal cortex, and the grey and white matter of the right planum temporale; brain regions implicated in both the language and the cognitive networks. The aPFC is part of the cingulo-opercular network (Dosenbach et al, 2008; Stockert et al, 2020) and is involved in almost all aspects of cognition (Burgess & Wu, 2013). In aphasia research specifically, increased functional activation in the cingulo-opercular network following anomia therapy has been found to correlate with improvements in word retrieval (Nardo, Holland, Leff, Price & Crinion, 2017; Fridriksson et al, 2006), possibly due to an improved ability to inhibit incorrect naming attempts. Although, changes in this brain region were not found to correlate with improvements in naming in the iTalkBetter study, increased GM in the aPFC may reflect changes in

abilities such as inhibiting incorrect responses, awareness of competence, or attention (Fleming, Weil, Nagy, Dolan & Rees, 2010; Burgess & Wu, 2013). The PT is involved in language processing and structural changes here have been identified following both comprehension and production interventions in PWA (Fleming et al, 2020; Wan et al, 2014). The reason for structural change in this region in the current study could be explained by the role the PT plays in speech comprehension due to the phonemic cueing nature of the therapy. Conversely, the posterior part of the PT provides a link between auditory areas and motor regions to map sound to articulation, therefore, changes here could be interpreted as a response to intensive word retrieval practice.

Fleming and colleagues employed the exact longitudinal VBM method utilised in this study to investigate structural changes following speech comprehension therapy in PWA, however, this is the first study to use this method to examine changes following word finding therapy (Ashburner, 2013). A benefit of using this type of analysis, in addition to the robust, within-subject design, is that I was also able to explore changes over the pre-therapy period and therapy period, separately. This analysis revealed relatively stable grey matter volume in the left frontal lobe over the pre-therapy period, however, atrophy of both grey and white matter in the right temporal lobe which then curtailed over therapy. Brain atrophy in those who have suffered from a stroke has been found to be greater than expected in normal aging in a number of studies (Brodtmann et al, 2021; Seghier, Ramsden, Lim, Leff & Price, 2014). Due to this, understanding the trajectory of neurodegeneration in stroke and identifying interventions that can slow down, or reverse brain volume loss, is a particularly important research focus. The current results support the use of multiple brain scanning time-points to further our understanding of brain atrophy post-stroke,

and to investigate how this neurodegeneration may be impacted by behavioural treatments.

In **Chapter 3**, I further examined therapy-driven neuroplasticity mechanisms using functional task-based MRI imaging in the same subset of participants as above. In contrast to the structural imaging analysis, changes in functional activation were investigated by comparing the average of the two pre-therapy scans with the post-therapy scan, rather than analysing changes over the pre-therapy period to changes over the post-therapy period. As with the structural imaging, changes in task-related activation were not related to treatment-response, however, they were related to dose of therapy, suggesting the more therapy participants completed, the higher the increases in activity. This is the first study, to my knowledge, to investigate neuroplasticity in relation to treatment dose in a group of people with chronic post-stroke aphasia.

Changes in activation were found bilaterally in areas supporting speech processing (the primary auditory cortex) and speech articulation (the insula). Although both the left and the right PAC are involved in the early stages of auditory analysis, changes in this brain region have been found following word finding therapy, particularly for intervention programmes which have utilised phonological methods, as in the iTalkBetter therapy (Marcotte et al, 2018; Abel, Weiller, Huber, Willmes & Specht, 2015). One the other hand, the insula is thought to mediate articulation and plays a role in the co-ordination of higher cognitive aspects of the language network (Buchsbaum, Hickok & Humphries, 2001; Buchsbaum, Olsen, Koch & Berman, 2005; Hickok & Poeppel, 2015; Skipper-Kallal, Lace, Xing & Turkeltaub, 2017; Oh, Duerden & Pang, 2014). In naming therapy studies, changes in activation in the insula have been found predominantly in the right hemisphere (Nardo, Holland, Leff,

Price & Crinion, 2017; Raboyeau et al, 2008; Richter, Miltner & Straub, 2008). Although in the current study these changes were found bilaterally, similar interpretations, such as the role the insulae play in speech articulation, speech perception and executive functioning, could explain the increased activation here following iTalkBetter therapy.

Interestingly, the areas of change identified across the structural and functional imaging both include areas involved in speech perception and the cingulo-opercular network. Although there are differences between the two analyses in terms of exactly which areas were found to change, there was overlap for the cluster in the right hemisphere. This is perhaps not surprising considering the homologue region to this brain area in the left hemisphere was damaged for many of the participants who took part in the imaging section of this study. However, due to differences in the methods utilised to examine structural and functional changes, it is difficult to directly compare the two results. As changes were found in the right hemisphere over the pre-therapy period in the structural imaging, it may be that there were further functional changes here which were masked out by averaging the two pre-therapy scans in the functional imaging analysis. Therefore, a future suggestion for the data collected in this study would be to examine functional activation over the pre-therapy period in comparison to the post-therapy period. This analysis may reveal differential results and may even provide more statistical power in identifying areas of change that do correlate with behavioural outcomes.

6.2 Limitations and future directions

Limitations

My thesis investigated the efficacy of the iTalkBetter therapy in improving word retrieval in a single word naming test and in a spoken picture description task. However, the impact of therapy on participants' real word, functional communication was not explored. Although the semi-structured interviews which were completed before and after intervention did question participants and their conversation partners about whether taking part in the study impacted everyday life, these interviews were not analysed in this thesis. Furthermore, these were not an in-depth exploration of communication changes, but, instead, were to gain an overall insight of individuals' experience of the research process. As significant improvements were found in the spoken picture description task, an important consideration for future research would be to assess changes in word retrieval across different communicative contexts to fully understand the efficacy of iTalkBetter in improving word finding abilities (Webster, Whitworth & Morris, 2015; Herbert, Best, Hickin, Howard & Osbourne, 2003; Conroy, Sage and Lambon Ralph, 2009c). Despite gains made following impairment-based therapies in aphasia, people often continue to face challenges in real-word communication (Carragher, Conroy, Sage & Wilkinson, 2012). It is, therefore, vital that researchers and clinicians examine whether the intervention programmes they are providing lead to functional improvements.

Another key limitation is that the structural and functional neuroplasticity observed was not related to treatment response. The structural imaging results were associated with taking part in therapy, suggesting a more binary effect which is perhaps a form of 'priming of the system' or 'readying of the system'. On the other hand, the functional results were related to dose of therapy, suggesting increased task-related activity changes were monotonically linked to practice. The fact that neither were related to behavioural change may be due to the small sample size

included, or because of the heterogeneity between individuals in the strategies (and, therefore, brain regions) employed to support recovery (Kiran & Thompson, 2019). Additionally, for the functional imaging, activity changes may not have been associated with changes in single word retrieval because the tasks completed within the scanner were not directly related to the iTalkBetter therapy.

Although the current results do not provide a clear picture of the neural basis of recovery for this set of participants for this intervention programme, they do provide evidence of neuroplasticity in chronic stroke, furthering the evidence-base supporting the implementation of therapy programmes even many years after a stroke occurs. To explore whether there were individual differences in neuroplasticity for both the structural and functional data which explain behavioural change, a future consideration would be to use a multivariate analysis approach to analyse the current data.

Future directions

Due to the positive results found in this study, the iTalkBetter app will now be 'rolled-out' across both Android and Apple devices, allowing PWA across the country, and indeed the world, access to this intervention programme. This will enable those with aphasia to complete therapy independently and achieve the high-dose of therapy required to improve language outcomes. This efficacious and cost-effective method of delivering single word, anomia intervention will provide individuals and clinicians with a tool which is feasible to use within the current rehabilitation model.

Rolling-out the therapy will also greatly benefit the current research by enabling the implementation and testing of different intervention paradigms to see how iTalkBetter can be improved and individualised for each PWA. For example, a key question is

what dose and intensity schedule is optimal for delivering therapy. It is likely that person-related factors (such as severity, age and time since stroke) play a role in much and how frequent therapy needs to improve language outcomes (Brady et al, 2022). Hopefully, by making iTalkBetter available online, a large amount of data can be collected from many individuals so more in-depth analyses of dose-behavioural relationships can be investigated. Furthermore, it also presents an opportunity to conduct adaptive trial designs using dose-response modelling to identify optimal dose and intensity schedules on an individual basis (Doogan, Dignam, Copland & Leff, 2018; Bhatt & Mehta, 2016).

In the roll-out, baseline testing and demographic questionnaires will also be completed by anybody who uses the app. By using this data, future research could examine, with a much larger sample of participants, associations between person-related factors and improvements to ascertain if response to iTalkBetter can be predicted. Using this sort of modelling, PWA could be given estimates of how much therapy they would need to carry-out in order to improve, or perhaps be directed to therapy programmes more suitable for them. It may even be possible to collect clinical brain scans from those using the therapy app to combine brain and behavioural data to examine multiple factors which may explain and/ or predict improvements (Aguilar et al, 2018).

A further consideration for the iTalkBetter therapy would be to enable participants to choose the words they practice in therapy. In the clinical trial, although the word list chosen included items PWA had said they wanted to learn in a previous study, the word list provided was not individualised for the participants in the current study. It is imperative that researchers and clinicians work with PWA to choose words which will maximise the impact of intervention and have a functional benefit in everyday life

(Palmer, Chater and Hughes). This is especially important in iTalkBetter as, like many other impairment-based interventions, improvements at the single word level did not generalise to untrained items. One way in which the therapy could be personalised, that would relatively easy to implement at such a large scale in the roll-out, would be to have word lists relating to different aspects of life and communication (for example, ‘in the home’) which people could choose from. Those words could then be practiced until they are considered ‘achieved’ before moving on to a new set of words. This would not only give PWA autonomy over what they learn in therapy, and hopefully make iTalkBetter more functionally beneficial, but also provide a more ongoing therapy programme.

7 References

Abel, S., Schultz, A., Radermacher, I., Willmes, K. & Huber, W. (2005). Decreasing and increasing cues in naming therapy for aphasia. *Aphasiology*, *19*(9): 831–848.

Abel, S., Weiller, C., Huber, W., Willmes, K., & Specht, K. (2015). Therapy-induced brain reorganization patterns in aphasia. *Brain (London, England: 1878)*, *138*(4), 1097–1112.

Agostini, M., Garzon, M., Benavides-Varela, S., De Pellegrin, S., Bencini, G., Rossi, G…… Tonin, P. (2014). Telerehabilitation in poststroke anomia. *BioMed Research International*, 1-6.

Aguilar, O. M., Kerry, S. J., Ong, Y.-H., Callaghan, M. F., Crinion, J., Woodhead, Z. V. J.… Hope, T. M. H. (2018). Lesion-site-dependent responses to therapy after aphasic stroke. *Journal of Neurology, Neurosurgery & Psychiatry, 89*(12), 1352–1354.

Allendorfer, J. B., Kissela, B. M., Holland, S. K., & Szaflarski, J. P. (2012). Different patterns of language activation in post-stroke aphasia are detected by overt and covert versions of the verb generation fMRI task. *Medical Science Monitor*, *18*(3), CR135–CR147.

Ansaldo, A. I., & Saidi, L. G. (2014). Aphasia therapy in the age of globalization: Cross-linguistic therapy effects in bilingual aphasia. *Behavioural Neurology,* 603085-603085.

Armstrong, E. (2005). Expressing opinions and feelings in aphasia: Linguistic options. *Aphasiology*, *19*, 285–295.

Armstrong, E., & Ulatowska, H. (2007). Making stories: Evaluative language and the aphasia experience. *Aphasiology*, *21*, 763–774.

Ashburner, J. (2013). Symmetric diffeomorphic modeling of longitudinal structural MRI. *Frontiers in Neuroscience, 6*, 197.

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry - The methods. *NeuroImage, 11*(6), 805–821.

Awad, M., Warren, J. E., Scott, S. K., Turkheimer, F. E., & Wise, R. J. S. (2007). A common system for the comprehension and production of narrative speech. *The Journal of Neuroscience*, *27*(43), 11455–11464.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Baddeley, A., & Wilson, B. A. (1994). When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia*, *32*, 53-68.

Bangert, M., & Schlaug, G. (2006). Specialization of the specialized in features of external human brain morphology. *Eur. J. Neurosci*. *24*, 1832–1834.

Barbeau, E. B., Chai, X. J., Chen, J.-K., Soles, J., Berken, J., Baum, S., Watkins, K. E., & Klein, D. (2017). The role of the left inferior parietal lobule in second language learning: An intensive language training fMRI study. *Neuropsychologia*, *98*, 169–176.

Barbera, D.S., Huckvale, M., Fleming, V., Upton, E., Coley-Fisher, H., Doogan, C…… Crinion, J. (2021). NUVA: A Naming Utterance Verifier for Aphasia Treatment. *Computer Speech & Language*, *69*, 101221.

Barrows, P. D., & Thomas, S. A. (2017). Assessment of mood in aphasia following stroke: psychometric properties of the Dynamic Visual Analogue Mood Scales (D-VAMS). *Clinical Rehabilitation*, *32*(1), 94-102.

Basso, A. (2005). How intensive/prolonged should an intensive/prolonged treatment be? *Aphasiology, 19*(10-11), 975-984.

Beeke, S., Maxim, J. E., & Wilkinson, R. (2007). Using conversation analysis to assess and treat people with aphasia. *Seminars in Speech and Language*, *28*, 136–147.

Beglinger, L. J., Unverzagt, F. W., Beristain, X., & Kareken, D. (2008). An updated version of the Weigl discriminates adults with dementia from those with mild impairment and healthy controls. *Arch Clin Neuropsychol, 23*(2), 149-156.

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

Bezzola, L., Merillat, S., Gaser, C., & Jancke, L. (2011). Training-induced neural plasticity in golf novices. *J. Neurosci., 31*, 12444–12448.

Bhatt, & Mehta, C. (2016). Adaptive Designs for Clinical Trials. *The New England Journal of Medicine*, *375*(1), 65–74.

Bhogal, S. K., Teasell, R., & Speechley, M. (2003). Intensity of aphasia therapy: Impact on recovery. *Stroke: Journal of the American Heart Association, 34*(4), 987-992.

Binney, R., Zuckerman, B., & Reilly, J. (2016). A neuropsychological perspective on abstract word representation: From theory to treatment of acquired language disorders. *Current Neurology and Neuroscience Reports, 16*(9), 1-8.

Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech Language Pathology, 13*(3), 236-249.

Boyle, M. (2010). Semantic feature analysis treatment for aphasic word retrieval impairments: What's in a name? *Topics in Stroke Rehabilitation, 17*(6), 411-422.

Brady, M. C., Ali, M., Vandenberg, K., Williams, L., Williams, L. R., Abo, M……Harris Wright, H. (2022). Precision rehabilitation for aphasia by patient age, sex, aphasia severity, and time since stroke? A prespecified, systematic review-based, individual participant data, network, subgroup meta-analysis. *International Journal of Stroke,* (In press).

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke. *The Cochrane database of systematic reviews (6).*

Breitenstein, C., Grewe, T., Flöel, A., Ziegler, W., Springer, L., Martus, P., . . . Bamborschke, S. (2017). Intensive speech and language therapy in patients with chronic aphasia

after stroke: A randomised, open-label, blinded-endpoint, controlled trial in a health-care setting. *The Lancet, 389*(10078), 1528-1538.

Broca, P. (1861a). Nouvelle observation d'aphémie produite par une lésion de la troisième circonvolution frontale. *Bulletins de La Société d'anatomie (Paris), 6*(2e serie), 398–407.

Broca, P. (1861b). Perte de la parole: ramollissement chronique et destruction partielle du lobe anterieur gauche du cerveau. *Bulletins de La Societe d'anthropologie, 2*, 2(1re serie), 235–238.

Broca, P. (1864). Sur les mots aphemie, aphasie et aphrasie. *Lettre a M. Le Professeur Trousseau, Gazette Des Hopitaux, 23*.

Brodtmann, A., Khlif M. S., Egorova, N., Veldsman, M., Bird, L. J., & Werden, E. (2020). Dynamic regional brain atrophy rates in the first year after ischemic stroke. *Stroke, 51*, e183–92.

Brodtmann, A., Werden, E., Khlif, M. S., Bird, L. J., Egorova, N., Veldsman, M…… Donnan, G. (2021). Neurodegeneration over 3 years following ischaemic stroke: Findings from the cognition and neocortical volume after stroke study. *Frontiers in Neurology*, *12*, 754204–754204.

Brownsett, S., Warren, J. E., Geranmayeh, F., Woodhead, Z., Leech, R., & Wise, R. J. S. (2014). Cognitive control and its impact on recovery from aphasic stroke. *Brain*, *137*(1), 242–254.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2013). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911.

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990.

Buchsbaum, B. R., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*, *25*(5), 663–678.

Buchsbaum, B. R., Olsen, R. K., Koch, P., & Berman, K. F. (2005). Human Dorsal and Ventral Auditory Streams Subserve Rehearsal-Based and Echoic Processes during Verbal Working Memory. *Neuron*, *48*(4), 687–697.

Burgess, P., & Wu, H. C. (2013). Rostral Prefrontal Cortex (Brodmann Area 10). In D. T. Stuss and R. T. Knight (Eds). *Principles of Frontal Lobe Function*. England: Oxford University Press.

Burgess, P., Dumontheil, I., & Gilbert, S. J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Sciences*, *11*(7), 290–298.

Buxton, R. B. (2002). *Introduction to Functional Magnetic Resonance Imaging: Principles and Techniques*. Cambridge, UK: Cambridge University Press.

Cao, Y., Vikingstad, E. M., George, K. P., Johnson, A. F., & Welch, K. M. A. (1999). Cortical language activation in stroke patients recovering from aphasia with functional MRI. *Stroke*, *30*(11), 2331–2340.

Carragher, M., Conroy, P., Sage, K., & Wilkinson, R. (2012). Can impairment-focused therapy change the everyday conversations of people with aphasia? A review of the literature and future directions. *Aphasiology, 26*(7), 895-916.

Catani, M., Allin, Matthew, P. G., Husain, M., Pugliese, L., Mesulam, M. M., Murray, R. M., & Jones, D. K. (2007). Symmetries in human brain language pathways correlate with verbal recall. *Proceedings of the National Academy of Sciences, 104*(43), 17163–17168.

Cattell, R., & Cattell, K. (1949). *Culture Fair Intelligence Tests*. Champaign, Illinois, U.S.A.: Institute for personality and ability testing.

Chang, A. J., Wilmskoetter, J., Fridriksson, J., McKinnon, E. T., Johnson, L. P., Basilakos, A., Jensen, J. H., Rorden, C., & Bonilha, L. (2021). Cortical microstructural changes associated with treated aphasia recovery. *Annals of Clinical and Translational Neurology*, *8*(9), 1884–1894.

Cherney, L. R., Patterson, J. P., & Raymer, A. M. (2011). Intensity of aphasia therapy: Evidence and efficacy*. Current Neurology and Neuroscience Reports, 11*(6), 560–569.

Clarke, D. J., Burton, L., Tyson, S. F., Rodgers, H., Drummond, A., Palmer, R……Forster, A. (2018). Why do stroke survivors not receive recommended amounts of active therapy? Findings from the ReAcT study, a mixed-methods case-study evaluation in eight stroke units. *Clinical Rehabilitation*, *32*(8), 1119-1132.

Cocquyt, E. M., De Ley, L., Santens, P., Van Borsel, J., & De Letter, M. (2017). The role of the right hemisphere in the recovery of stroke-related aphasia: A systematic review. *Journal of Neurolinguistics*, *44*, 68–90.

Code, C. & Heron, C. (2003). Services for aphasia, other acquired adult neurogenic communication and swallowing disorders in the United Kingdom, 2000. *Disability Rehabilitation*, *25(21),* 1231-7.

Code, C., Torney, A., Gildea-Howardine, E., & Willmes, K. (2010). Outcome of a one-month therapy intensive for chronic aphasia: Variable individual responses. *Seminars in Speech and Language, 31*, 21-33.

Conroy, P., & Scowcroft, J. (2012). Decreasing cues for a dynamic list of noun and verb naming targets: A case-series aphasia therapy study. *Neuropsychological Rehabilitation*, *22*(2), 295–318.

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009a). Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology*, *23*(11), 1311–1337.

Conroy, P., Sage, K., & Lambon Ralph, M. A. (2009b). The effects of decreasing and increasing cue therapy on improving naming speed and accuracy for verbs and nouns in aphasia. *Aphasiology*, *23*(6), 707–730.

Conroy, P., Sage, K., & Lambon Ralph, M. (2009c). Improved vocabulary production after naming therapy in aphasia: Can gains in picture naming generalise to connected speech? *International Journal of Language and Communication Disorders, 44*(6), 1036-1062.

Conroy, P., Snell, C., Sage, K. E., & Lambon Ralph, M. A. (2012). Using Phonemic Cueing of Spontaneous Naming to Predict Item Responsiveness to Therapy for Anomia in Aphasia. *Archives of Physical Medicine and Rehabilitation*, *93*(1), S53–S60.

Conroy, P., Sotiropoulou Drosopoulou, C., Humphreys, G. F., Halai, A. D., & Lambon Ralph, M. A. (2018). Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia. *Brain*, *141*(6), 1815–1827.

Cordella, C., Munsell, M., Godlove, J., Anantha, V., Advani, M., Kiran, S. (2022). Dosage frequency effects on treatment outcomes following self-managed digital therapy: Retrospective cohort study. *Journal of Medical Internet Research, 24*(7), e36135.

Crinion, J. T., & Leff, A. P. (2015). Using functional imaging to understand therapeutic effects in poststroke aphasia. *Current Opinion in Neurology, 28*(4), 330-37.

Crinion, J. T., & Price, C. J. (2005). Right anterior superior temporal activation predicts auditory sentence comprehension following aphasic stroke. *Brain (London, England : 1878)*, *128*(12), 2858–2871.

Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain*, *128*(3), 615–627.

Culton, G. L. (1969). Spontaneous recovery from aphasia. *Journal of Speech, Language and Hearing Research, 12*, 825–32.

Davis, A., & Pring, T. (1991). Therapy for word-finding deficits: More on the effects of semantic and phonological approaches to treatment with dysphasic patients. *Neuropsychological Rehabilitation*, *1,* 135–145.

De Smet, H. J., Paquier, P., Verhoeven, J., & Mariën, P. (2013). The cerebellum: Its role in language and related cognitive and affective functions. *Brain Lang, 127*(3), *334-42.*

de Zubicaray, G. I., Zelaya, F. O., Andrew, C., Williams, S. C., & Bullmore, E. (2000). Cerebral regions associated with verbal response initiation, suppression and strategy use. *Neuropsychologia*, *38*(9), 1292–1304.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., Gagnon, D.A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801-838.

Des Roches, C. A., & Kiran, S. (2017). Technology-based rehabilitation to improve communication after acquired brain injury. *Frontiers in Neuroscience, 11*, 382.

Dignam, Copland, D., O'Brien, K., Burfein, P., Khan, A., & Rodriguez, A. D. (2017). Influence of cognitive ability on therapy outcomes for anomia in adults with chronic post-stroke aphasia. *Journal of Speech, Language, and Hearing Research*, *60*(2), 406–421.

Dignam, J., Copland, D., McKinnon, E., Burfein, P., O'Brien, K., Farrell, A., & Rodriguez, A. D. (2015). Intensive versus distributed aphasia therapy: A non-randomised, parallel-group, dosage-controlled study. *Stroke*, *46*(8), 2206-2211.

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

Dignam, J., Rodriguez, A. D., & Copland, D. A. (2016). Evidence for intensive aphasia therapy: Consideration of theories from neuroscience and cognitive psychology. *PM & R: The Journal of Injury, Function, and Rehabilitation, 8*(3), 254–267.

Doogan, C., Dignam, J., Copland, D., & Leff, A. (2018). Aphasia recovery: When, how and who to treat? *Current Neurology and Neuroscience Reports, 18*(12), 90.

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

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: Changes in grey matter induced by training. *Nature, 427*(6972), 311–312.

El Hachioui, H., Lingsma, H. F., Van De Sandt-Koenderman, M. W. M. E., Dippel, D. W. J., Koudstaal, P. J., & Visch-Brink, E. G. (2013). Long-term prognosis of aphasia after stroke. *Journal of Neurology, Neurosurgery and Psychiatry, 84*(3), 310-315.

Feigin, V. L., Norrving, B. and Mensah, G. A. (2017). Global Burden of Stroke. *Circulation Research*, *120*(3), 439-448.

Fernández Coello, A., Moritz-Gasser, S., Martino, J., Martinoni, M., Matsuda, R., & Duffau, H. (2013). Selection of intraoperative tasks for awake mapping based on relationships between tumor location and functional networks. *Journal of Neurosurgery, 119*(6), 1380-1394.

Field, A. (2013). *Discovering Statistics using IBM SPSS statistics; and sec and drugs and rock "n" roll (4th ed.)*. Portland, USA: Sage Publications.

Fillingham, J. K., Hodgson, C., Sage, K., and Lambon, M. A. (2003). The application of errorless learning to aphasic disorder: A review of theory and practice. *Neuropsychological Rehabilitation, 13*(3), 337-363.

Fillingham, J. K., Sage, K. and Lambon Ralph, M. A. 2005a. The treatment of anomia using errorless vs. errrorful learning: Are frontal executive skills and feedback important? *International Journal of Language and Communication Disorders*, *40*(4), 505–524.

Fillingham, J. K., Sage, K., & Lambon Ralph, M. A. (2005b). Further explorations and an overview of errorless and errorful therapy for anomia: The number of naming attempts during therapy effects outcome. *Aphasiology*, *19*(7), 597–614.

Fillingham, J. K., Sage, K., and Lambon Ralph M. A. (2006). The treatment of anomia using errorless learning. *Neuropsychological Rehabilitation, 16*(2), 129-154.

Fleming, V., Brownsett, S., Krason, A., Maegli, M. A., Coley-Fisher, H., Ong, Y.-H…… Leff, A. P. (2020). Efficacy of spoken word comprehension therapy in patients with chronic aphasia: a cross-over randomised controlled trial with structural imaging. *Journal of Neurology, Neurosurgery and Psychiatry*, *92*(4), 418–424.

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science, 329*(5998), 1541–1543.

Foygel, D., & Dell, G. S. (2000). Models of Impaired Lexical Access in Speech Production. *Journal of Memory and Language*, *43*(2), 182–216.

Fridriksson, J., Baker, J. M., & Moser, D. (2009). Cortical mapping of naming errors in aphasia. *Human Brain Mapping*, *30*(8), 2487–2498.

Fridriksson, J., Bonilha, L., Baker, J. M., Moser, D., & Rorden, C. (2010). Activity in preserved left hemisphere regions predicts anomia severity in aphasia. *Cerebral Cortex, 20*(5), 1013– 1019.

Fridriksson, J., den Ouden, D.-B., Hillis, A. E., Hickok, G., Rorden, C., Basilakos, A., … Bonilha, L. (2018). Anatomy of aphasia revisited. *Brain*, *141*(3), 848–862.

Fridriksson, J., Morrow-Odom, L., Moser, D., Fridriksson, A., & Baylis, G. (2006). Neural recruitment associated with anomia treatment in aphasia. *NeuroImage*, *32*(3), 1403–1412.

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences, 16*(5), 262–268.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. P., Frith, C. D, & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping, 2*(4), 189-210.

Fujii, M., Maesawa, S., Ishiai, S., Iwami, K., Futamura, M., & Saito, K. (2016). Neural basis of language: An overview of an evolving model. *Neurologia Medico-chirurgica, 56*(7), 379-86.

Geranmayeh, F., Brownsett, S. L. E., Leech, R., Beckmann, C. F., Woodhead, Z., & Wise, R. J. S. (2012). The contribution of the inferior parietal cortex to spoken language production. *Brain and Language*, *121*(1), 47–57.

Geranmayeh, F., Brownsett, S. L. E., & Wise, R. J. S. (2014). Task-induced brain activity in aphasic stroke patients: What is driving recovery? *Brain, 137*(10), 2632-648.

Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D., & Burgess, P. W. (2006). Functional specialization within rostral prefrontal cortex (Area 10): A meta-analysis. *Journal of Cognitive Neuroscience*, *18*(6), 932–948.

Godecke, E., Rai, T., Ciccone, N., Armstrong, E., Granger, A., & Hankey, G. (2013). Amount of therapy matters in very early aphasia rehabilitation after stroke: A clinical prognostic model. *Seminars in Speech and Language, 34*(3), 129-141.

Godlove, J., Anantha, V., Advani, M., Des Roches, C., & Kiran, S. (2019). Comparison of therapy practice at home and in the clinic: A retrospective analysis of the constant therapy platform data set. *Frontiers in Neurology, 25*, 10-140.

Golestani, N., Price, C. J., & Scott, S. K. (2011). Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *Journal of Neuroscience, 31*(11), 4213–4220.

Goodglass, H., & Wingfield, A. (1997). *Anomia: Neuroanatomical and Cognitive Correlates*. San Diego, CA: Academic Press.

Greenwood, A., Grassly, J., Hickin, J., & Best, W. (2010). Phonological and orthographic cueing therapy: A case of generalised improvement. *Aphasiology, 24*(9), 991-1016.

Griffis, J. S., Nenert, R., Allendorfer, J. B., Vannest, J., Holland, S., Dietz, A., & Szaflarski, J. P. (2017). The canonical semantic network supports residual language function in chronic post-stroke aphasia. *Human Brain Mapping*, *38*(3), 1636–1658.

Halai, A. D., Woollams, A. M., & Lambon Ralph, M. A. (2018). Triangulation of language-cognitive impairments, naming errors and their neural bases post-stroke. *NeuroImage: Clinical*, *17*, 465–473.

Haller, S., Radue, E. W., Erb, M., Grodd, W., & Kircher, T. (2005). Overt sentence production in event-related fMRI. *Neuropsychologia*, *43*(5), 807–814.

Harnish, S. M., Morgan, J., Lundine, J. P., Bauer, A., Singletary, F., & Benjamin, M. L. (2014). Dosing of a cued picture-naming treatment for anomia. *American Journal of Speech Language Pathology, 23*(2), S285–S299.

Harnish, S. M., Schwen Blackett, D., Zezinka, A., Lundine, J. P., & Pan, X. (2018). Influence of working memory on stimulus generalization in anomia treatment: A pilot study. *Journal of Neurolinguistics, 48*, 142-156.

Hartwigsen, G., & Saur, D. (2019). Neuroimaging of stroke recovery from aphasia – Insights into plasticity of the human language network. *NeuroImage*, *190*, 14–31.

Harvey, S., Carragher, M., Dickey, M. W., Pierce, J. E., & Rose, M. L. (2022). Dose effects in behavioural treatment of post-stroke aphasia: A systematic review and meta-analysis. *Disability and Rehabilitation*, *44*(12), 2548–2559.

Herbert, R., Best, W., Hickin, J., Howard, D., & Osborne, F. (2003). Combining lexical and interactional approaches to therapy for word finding deficits in aphasia. *Aphasiology, 17*(12), 1163-1186.

Herbert, R., Hickin, J., Howard, D., Osborne, F., & Best, W. (2008). Do picture-naming tests provide a valid assessment of lexical retrieval in conversation in aphasia? *Aphasiology, 22*, 184– 203.

Herdener, M., Esposito, F., di Salle, F., Boller, C., Hilti, C. C., Habermeyer, B……Cattapan-Ludewig, K. (2010). Musical training induces functional plasticity in human hippocampus. *The Journal of Neuroscience, 30*, 1377–84.

Hervais-Adelman, A., Moser-Mercer, B., Murray, M. M., & Golestani, N. (2017). Cortical thickness increases after simultaneous interpretation training. *Neuropsychologia, 98*, 212–219.

Hickin, J., Best, W., Herbert, R., Howard, D., & Osborne, F. (2002). Phonological therapy for word-finding difficulties: A re-evaluation. *Aphasiology, 16*(10-11), 981-999.

Hickok, G. (2009). The functional neuroanatomy of language. *Physics of Life Reviews*, *6*(3), 121–143.

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. *Nature Reviews Neuroscience*, *8*(5), 393–402.

Hickok, G., & Poeppel, D. (2015). Neural basis of speech perception. *Handbook of Clinical Neurology, 129*, 149-60.

Hillis, A. E. (1998). Treatment of naming disorders: New issues regarding old therapies. *Journal of the International Neuropsychological Society*, *4*, 648–660.

Holland, R., & Crinion, J. (2021). Cued naming and the left inferior frontal cortex. *Life Science Weekly.*

Holland, R., Johns, S., & Woollams, A. (2018). The impact of phonological versus semantic repetition training on generalisation in chronic stroke aphasia reflects differences in dorsal pathway connectivity. *Neuropsychological Rehabilitation, 28*(4), 548-567.

Hope, T., Leff, A. P., Prejawa, S., Bruce, R., Haigh, Z., Lim, L., Ramsden, S., Oberhuber, M., Ludersdorfer, P., Crinion, J., Seghier, M. L., & Price, C. J. (2017). Right hemisphere structural adaptation and changing language skills years after left hemisphere stroke. *Brain (London, England: 1878)*, *140*(6), 1718–1728.

Hope, T., Nardo, D., Holland, R., Ondobaka,S., Akkad, H., Price, C. J., Leff, A. P., & Crinion, J. (2021). Lesion site and therapy time predict responses to a therapy for anomia after stroke: A prognostic model development study. *Scientific Reports*, *11*(1), 18572–18572.

Hula, W. D., Panesar, S., Gravier, M. L., Yeh, F.-C., Dresang, H. C., Dickey, M. W., & Fernandez-Miranda, J. C. (2020). Structural white matter connectometry of word production in aphasia: An observational study. *Brain*, *143*(8), 2532–2544.

Iorga, M., Higging, J., Caplan, D., Zinbarg, R., Kiran, S., Thompson, C. K., Rapp, B., & Parrish, T. B. (2021). Predicting language recovery in post-stroke aphasia using behavior and functional MRI. *Scientific Reports, 11*(1), 8419–8419.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1), 101–144.

Johansson, M. B., Carlsson, M., & Sonnander, K. (2012). Communication difficulties and the use of communication strategies: from the perspective of individuals with aphasia. *International Journal of Language and Communication Disorders, 47*, 144–55.

Johnson, J. P., Meier, E. L., Pan, Y., & Kiran, S. (2019). Treatment-related changes in neural activation vary according to treatment response and extent of spared tissue in patients with chronic aphasia. *Cortex*, *121*, 147–168.

Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic Assessments of Language Processing in Aphasia: PALPA.* Hove, UK: Lawrence Erlbaum.

Kiran, S. (2016). How does severity of aphasia influence individual responsiveness to rehabilitation? Using big data to understand theories of aphasia rehabilitation. *Seminars in Speech and Language, 37*, 48-59.

Kiran, S., Meier, E. L., & Johnson, J. P. (2019). Neuroplasticity in aphasia: A proposed framework of language recovery. *Journal of Speech, Language, and Hearing Research*, *62*(11), 3973–3985.

Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of language networks in aphasia: Advances, updates, and future challenges. *Frontiers in neurology, 10*, 295–295.

Kiran, S., Sandberg, C., & Abbott, K. (2009). Treatment for lexical retrieval using abstract and concrete words in persons with aphasia: Effect of complexity. *Aphasiology*, *23*(7-8), 835–853.

Kirmess, M., & Maher, L. M. (2010). Constraint induced language therapy in early aphasia rehabilitation. *Aphasiology, 4,* 725-736.

Kleim, J. (2011). Neural plasticity and neurorehabilitation: Teaching the new brain old tricks. *Journal of Communication Disorders, 44*, 521 – 528.

Kleim, J., & Jones, T. A. (2008). Principles of experience-dependent neural plasticity: Implications for rehabilitation after brain damage. *Journal of Speech, Language, and Hearing Research*, *51*(1), S225–S239.

Knecht, S., Deppe. M., Dräger. B., Bobe. L., Lohmann. H., Ringelstein. E., & Henningsen. H. (2000). Language lateralization in healthy right-handers. *Brain, 123*, 74-81.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990.

Laine, M., & Martin, N. (2006). *Anomia: Theoretical and Clinical Aspects.* Hove: Psychology.

Lam, J. M., & Wodchis, W. P. (2010). The relationship of 60 disease diagnoses and 15 conditions to preference-based health-related quality of life in Ontario hospital-based long-term care residents. *Medical Care, 48*(4), 380-387.

Lambon Ralph, M. A., Snell, C., Fillingham, J. K., 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. *Neuropsychological Rehabilitation, 20*, 289-305.

Laska, A. C., Hellblom, A., Murray, V., Kahan, T., & Von Arbin, M. (2001). Aphasia in acute stroke and relation to outcome. *Journal of Internal Medicine, 249*(5), 413-422.

Leff, A. P., Crinion, J., Scott, S., Turkheimer, F., Howard, D., & Wise, R. (2002). A physiological change in the homotopic cortex following left posterior temporal lobe infarction. *Annals of Neurology, 51*(5), 553–558.

Leonard, E., Rochon, E., & Laird, L. (2008). Treating naming impairments in aphasia: Findings from a phonological components analysis treatment. *Aphasiology, 22*(9), 923-947.

Li, E. C., & Williams, S. E. (1991). An investigation of naming errors following semantic and phonemic cueing. *Neuropsychologia, 29*, 1083-1093.

Li, R., Mukadam, N., & Kiran, S. (2022). Functional MRI evidence for reorganization of language networks after stroke. *Handbook of Clinical Neurology*, *185*, 131–150.

Lorca-Puls, D. L., Gajardo-Vidal, A., White, J., Seghier, M. L., Leff, A. P., Green, D. W., … Price, C. J. (2018). The impact of sample size on the reproducibility of voxel-based lesion-deficit mappings. *Neuropsychologia, 115*, 101–111.

Lorenz, A., & Ziegler, W. (2009). Semantic vs. word form specific techniques in anomia treatment: A multiple single case study. *Journal of Neurolinguistics, 22*, 515-537.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage, 19*(3), 1233–1239.

Mandelli, M. L., Vilaplana, E., Brown, J. A., Hubbard, H. I., Binney, R. J., Attygalle, S…..… GornoTempini, M. L. (2016). Healthy brain connectivity predicts atrophy progression in nonfluent variant of primary progressive aphasia. *Brain, 139*(10), 2778–2791.

Manly, T., Davison, B., Heutink, J., Galloway, M., & Robertson, I. (2000). Not enough time or not enough attention? Speed, error and self-maintained control in the Sustained Attention to Response Test (SART). *Clinical Neuropsychological Assessment, 3*, 167-177.

Marcotte, K., & Ansaldo, A. I. (2010). The Neural Correlates of Semantic Feature Analysis in Chronic Aphasia: Discordant Patterns According to the Etiology. *Seminars in Speech and Language*, *17*(1), 052–063.

Marcotte, K., Daniel, A.-R., Damien, B., de Préaumont, M., Généreux, S., Hubert, M., & Ansaldo, A. (2012). Therapy-induced neuroplasticity in chronic aphasia. *Neuropsychologia, 50*(8), 1776–1786.

Marcotte, K., Laird, L., Bitan, T., Meltzer, J. A., Graham, S. J., Leonard, C., & Rochon, E. (2018). Therapy-induced neuroplasticity in chronic aphasia after phonological component analysis: A matter of intensity. *Frontiers in Neurology, 9*, 225.

Marshall, J., Devane, N., Edmonds, L., Talbot, R., Wilson, S., Woolf, C., & Zwart, N. (2018). Delivering word retrieval therapies for people with aphasia in a virtual communication environment. *Aphasiology, 32*(9), 1054-1074.

Mårtensson, J., Eriksson, J., Bodammer, N. C., Lindgren, M., Johansson, M., Nyberg, L., & Lövdén, M. (2012). Growth of language-related brain areas after foreign language learning. *NeuroImage, 63*(1), 240–244.

Martin, N., Fink, R. B., Renvall, K., & Laine, M. (2006). Effectiveness of contextual repetition priming treatments for anomia depends on intact access to semantics. *Journal of the International Neuropsychological Society*, *12*(6), 853-866.

Mason, C., Nickels, L., McDonald, B., Moses, M., Makin, K. & Taylor, C. (2011). Treatment of word retrieval impairments in aphasia: Evaluation of a self-administered home programme using personally chosen words. *Aphasiology, 25*(2), 245-268.

McKissock S., & Ward, J. (2007). Do errors matter? Errorless and errorful learning in anomic picture naming. *Neuropsychological Rehabilitation, 17*(3), 355-373.

Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., & Price, C. J. (2004). Structural plasticity in the bilingual brain. *Nature, 431*(7010), 757–757.

Mechelli, A., Josephs, O., Lambon Ralph, M. A., McClelland, J. L., Price, C. J. (2007). Dissociating stimulus-driven semantic and phonological effect during reading and naming. *Human Brain Mapping, 28*, 205-17.

Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2005). Voxel-Based Morphometry of the Human Brain: Methods and Applications. *Current Medical Imaging Reviews, 1*(1).

Meier, E. L., Kapse, K. J., & Kiran, S. (2016). The relationship between frontotemporal effective connectivity during picture naming, behavior, and preserved cortical tissue in chronic aphasia. *Frontiers in Human Neuroscience*, *10*(2016), 109–109.

Meinzer, M., Beeson, P. M., Cappa, S., Crinion, J., Kiran, S., Saur, D…….Thompson, C. K. (2013). Neuroimaging in aphasia treatment research: Consensus and practical guidelines for data analysis. *NeuroImage*, *73*, 215–224.

Meinzer, M., Elbert, T., Wienbruch, C., Djundja, D., Barthel, G., & Rockstroh, B. (2004). Intensive language training enhances brain plasticity in chronic aphasia. *BMC Biology, 2*(20).

Menke, R., Meinzer, M., Kugel, H., Deppe, M., Baumgärtner, A., Schiffbauer, H., … Breitenstein, C. (2009). Imaging short- and long-term training success in chronic aphasia. *BMC Neuroscience, 10*, 118.

Meteyard, L., & Bose, B. (2018). What does a cue do? Comparing phonological and semantic cues for picture naming in aphasia*. Journal of Speech, Language, and Hearing Research, 61*(3), 658-674.

Middleton, E. L., Schwarts, M. F., Rawson, K. A., Traut, H., & Verkuilen, J. (2016). Towards a theory of learning for naming rehabilitation: Retrieval practice and spacing effects. *Journal of Speech, Language, and Hearing Research, 59*(5), 1111-1122.

Mills, K. L., & Tamnes, C. K. (2014). Methods and considerations for longitudinal structural brain imaging analysis across development. *Developmental Cognitive Neuroscience, 9*, 172–190.

Morris, R., Eccles, A., Ryan, B., & Kneebone, I. (2017). Prevalence of anxiety in people with aphasia after stroke. *Aphasiology, 31*(12), 1410-1415.

Munsell, M., De Oliviera, E., Saxena, S., Godlove, J., & Kiran, S. (2020). Closing the digital divide in speech, language, and cognitive therapy: Cohort study of the factors associated with technology usage for rehabilitation. *Journal of medical Internet research, 22*(2), e16286–e16286.

Nardo, D., Holland, R., Leff, A. P., Price, C. J., & Crinion, J. T. (2017). Less is more: Neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain, 140*(11), 3039-3054.

Neumann, Y. (2018). A case series comparison of semantically focused vs. phonologically focused cued naming treatment in aphasia. *Clinical Linguistics and Phonetics, 32*(1), 127.

Newton, K., & Barry, C. (1997). Concreteness effects in word production but not word comprehension in deep dyslexia. *Cognitive Neuropsychology*, *14*(4), 481–509.

Nickels, L. (2002). Therapy for naming disorders: Revisiting, revising and reviewing. *Aphasiology, 16*(10-11), 935-979.

Nickels, L., & Best, W. (1996). Therapy for naming disorders (part I): Principles, puzzles and progress. *Aphasiology, 10*(1), 21-47.

Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia, 33*(10), 1281–1303.

O'Donnell, L. J., & Westin, C.-F. (2011). An introduction to diffusion tensor image analysis. *Neurosurgery Clinics of North America, 22*(2), 185–196.

Off, C. A., Griffin, J. R., Spencer, K. A., & Rogers, M. A. (2016). The impact of dose on naming accuracy with persons with aphasia. *Aphasiology, 30*(9), 983-1011.

Oh, A., Duerden, E. G., & Pang, E. W. (2014). The role of the insula in speech and language processing. *Brain and Language*, *135*, 96–103.

Okada, K., & Hickok, G. (2006). Identification of lexical–phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. *Neuroreport*, *17*(12), 1293–1296.

Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I.-H., Saberi, K., … Hickok, G. (2010). Hierarchical organization of human auditory cortex: Evidence from acoustic invariance in the response to intelligible speech. *Cerebral Cortex*, *20*(10), 2486–2495.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227–252.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. *Psychological Review, 56*, 132–143.

Palmer, R., Chater, T., & Hughes, H. (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*(3), E0174065.

Palmer, R., Dimairo, M., Cooper, C., Enderby, P., Brady, M., Bowen, A…..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. *The Lancet Neurology, 18*(9), 821-833.

Palmer, R., Enderby, P., Cooper, C., Latimer, N., Julious, S., Paterson, G., … Hughes, H. (2012). Computer therapy compared with usual care for people with long-standing aphasia poststroke: A pilot randomized controlled trial (step by step). *Stroke, 43*, 1904–1911.

Palmer, R., Witts, H. and Chater, T. (2018). What speech and language therapy do community dwelling stroke survivors with aphasia receive in the UK? *PLoS ONE*, *13*(7), e0200096.

Papathanasiou, I., Coppens, P., & Davidson, B. (2017). Aphasia and related neurogenic communication disorders: Basic concepts, management and efficacy. In I. Papathanasiou & P. Coppens (Eds.), *Aphasia and Related Neurogenic Communication Disorders* (2nd eds., pp. 3-10). Burlington, MA: Jones & Bartlett Learning.

Patterson, K.E., & Shewell, C. (1987). Speak and spell: Dissociations and word-class effects. In M. Coltheart, R. Job, & G. Sartori (Eds.). *The Cognitive Neuropsychology of Language*. Hove, UK: Lawrence Erlbaum Associates Inc.

Pearce, A. J., Thickbroom, G. W., Byrnes, M. L., & Mastaglia, F. L. (2000). Functional reorganisation of the corticomotor projection to the hand in skilled racquet players. *Exp. Brain Res, 130*, 238–243.

Penhune, V. B., Zatorre, R. J., MacDonald, J. D., & Evans, A. C. (1996). Interhemispheric anatomical differences in human primary auditory cortex: Probabilistic mapping and volume measurement from magnetic resonance scans. *Cerebral Cortex*, *6*(5), 661–672.

Perani, D., Cappa, S., Tettamanti, M., Rosa, M., Scifo, P., Miozzo, A., Basso, A., & Fazio, F. (2003). A fMRI study of word retrieval in aphasia. *Brain and Language*, *85*(3), 357–368.

Persad, C., Wozniak, L., & Kostopoulous, E. (2013). Retrospective analysis of outcomes from two intensive comprehensive aphasia programs. *Topics in Stroke Rehabilitation, 20*, 388-397.

Petrides, M., & Pandya, D. N. (2007). Efferent association pathways from the rostral prefrontal cortex in the macaque monkey. *The Journal of Neuroscience*, *27*(43), 11573–11586.

Plowman, E., Hentz, B., & Ellis, C. (2012). Post-stroke aphasia prognosis: A review of patientrelated and stroke-related factors. *Journal of Evaluation in Clinical Practice, 18*(3), 689694.

Pompon, R. H., Bislick, L., Elliott, K., Madden, E. B., Minkina, I., Oelke, M., & Kendall, D. (2017). Influence of linguistic and nonlinguistic variables on generalization and maintenance following phonomotor treatment for aphasia. *American Journal of Speech Language Pathology, 26*(4), 1092-1104.

Price, C. J. (1998). The functional anatomy of word comprehension and production. Trends in *Cognitive Science, 2*, 281-8.

Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences, 1191*, 62 – 88.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage, 62*(2), 816-847.

Price, C. J., & Crinion, J. (2005). The latest on functional imaging studies of aphasic stroke. *Current Opinion in Neurology, 18*(4), 429-434.

Price, C. J., Seghier, M. L., & Leff, A. P. (2010). Predicting language outcome and recovery after stroke: The PLORAS system. *Nature Reviews Neurology, 6*(4), 202-210.

Python, G., Pellet Cheneval, P., Bonnans, C., & Laganaro, M. (2021). Facilitating Word Retrieval in Aphasia: Which Type of Cues for Which Aphasic Speakers? *Frontiers in Human Neuroscience*, *15*, 747391–747391.

Raboyeau, G., De Boissezon, X., Marie, N., Balduyck, S., Puel, M., Bézy, C., Démonet, J. F., & Cardebat, D. (2008). Right hemisphere activation in recovery from aphasia: Lesion effect or function recruitment? *Neurology*, *70*(4), 290–298.

Ramsberger, M., & Marie, B. (2007). Self-administrated cued naming therapy: A single-participant investigation of a computer-based therapy program replicated in four cases. *American Journal of Speech-Language Pathology*, *16*, 343–358.

Rasmussen, T., & Milner, B. (1975). Clinical and surgical studies of the cerebral speech areas in man. In K. J. Zulch, O. Creutzfeld & G. C. Galbraith (Eds). *Cerebral localization* (pp.238-257). New York: Springer-Verlag.

Raymer, A. M., Beeson, P., Holland, A., Kendall, D., Maher, L. M., Martin, N., . . . Gonzalez Rothi, L. J. (2008). Translational research in aphasia: From neuroscience to neurorehabilitation. *Journal of Speech, Language, and Hearing Research*, *51*(1), S259-S275.

Renvall, K., & Nickels, L. (2019). Using treatment to improve the production of emotive adjectives in aphasia: A single-case study. *Aphasiology*, *33*(11), 1348–1371.

Renvall, K., Nickels, L., & Davidson, B. (2013). Functionally relevant items in the treatment of aphasia (part I): Challenges for current practice. *Aphasiology*, *27*(6), 636–650.

Richardson, J. (1976). Imageability and concreteness. *Bulletin of the Psychonomic Society*, *7*(5), 429–431.

Richter, M., Miltner, W. H. R., & Straube, T. (2008). Association between therapy outcome and right-hemispheric activation in chronic aphasia. *Brain*, *131*(5), 1391–1401.

Robey, R. R. (1998). A meta-analysis of clinical outcomes in the treatment of aphasia. *Journal of Speech, Language, and Hearing Research, 41*, 172-187.

Rochon, E., Leonard, C., Burianova, H., Laird, L., Soros, P., Graham, S., & Grady, C. (2010). Neural changes after phonological treatment for anomia: An fMRI study. *Brain and Language*, *114*(3), 164–179.

Rorden, C., & Brett, M. (2000). Stereotaxic display of brain lesions. *Behavioural Neurology, 12*(4), 191–200.

Rosen, H. J., Petersen, S. E., Linenweber, M. R., Snyder, A. Z., White, D. A., Chapman, L., Dromerick, A. W., Fiez, J. A., & Corbetta, M. (2000). Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology*, *55*(12), 1883–1894.

Routhier, S., Bier, N., & Macoir, J. (2016). Smart tablet for smart self-administered treatment of verb anomia: Two single-case studies in aphasia. *Aphasiology, 30*(2-3), 269-289.

Sage, K., Snell, C., & Lambon Ralph, M. A. (2011). How intensive does anomia therapy for people with aphasia need to be? *Neuropsychological Rehabilitation, 21*(1), 26-41.

Sandberg, C., & Kiran, S. (2014). How justice can affect jury: training abstract words promotes generalisation to concrete words in patients with aphasia. *Neuropsychological Rehabilitation, 24*(5), 738–69.

Sanjuán, A., Hope, T. M., Parker Jones, O., Prejawa, S., Oberhuber, M., Guerin, J., Seghier, M. L., Green, D. W., & Price, C. J. (2014). Dissociating the semantic function of two neighbouring subregions in the left lateral anterior temporal lobe*. Neuropsychologia, 76*, 153-162.

Santos-Santos, M. A., Mandelli, M. L., Binney, R. J., Ogar, J., Wilson, S. M., Henry, M. L.,…….Gorno-Tempini, M. L. (2016). Features of patients with nonfluent/agrammatic

primary progressive aphasia with underlying progressive supranuclear palsy pathology or corticobasal degeneration. *JAMA Neurology, 73*(6), 733.

Sarno, M. T., & Levita, E. (1971). Natural course of recovery in severe aphasia. *Arch Phys Med Rehabil, 52*, 175–8.

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain*, *129*(6), 1371–1384.

Saur, D., Schelter, B., Schnell, S., Kratochvil, D., Küpper, H., Kellmeyer, P., … Weiller, C. (2010). Combining functional and anatomical connectivity reveals brain networks for auditory language comprehension. *NeuroImage, 49*(4), 3187–3197.

Schacter, D. L., Dobbins, I. G., & Schnyer, D. M. (2004). Specificity of priming: A cognitive neuroscience perspective. *Nature Reviews Neuroscience, 5*(11), 853-862.

Schlaug, G., Marchina, S., & Norton, A. (2009). Evidence for plasticity in white-matter tracts of patients with chronic broca's aphasia undergoing intense intonation-based speech therapy. *Annals of the New York Academy of Sciences, 1169*, 385–394.

Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Faseyitan, O., Brecher, A., Dell, G. S., & Coslett, H. B. (2009). Anterior temporal involvement in semantic word retrieval: Voxel-based lesion-symptom mapping evidence from aphasia. *Brain*, *132*(12), 3411–3427.

Seeley, W. W. (2019). The salience network: A neural system for perceiving and responding to homeostatic demands. *The Journal of Neuroscience*, *39*(50), 9878–9882.

Seghier, M. L., Ramlackhansingh, A., Crinion, J., Leff, A. P., & Price, C. J. (2008). Lesion identification using unified segmentation-normalisation models and fuzzy clustering. *NeuroImage, 41*(4), 1253–1266.

Seghier, M. L., Ramsden, S., Lim, L., Leff, A. P., & Price, C. J. (2014). Gradual lesion expansion and brain shrinkage years after stroke. *Stroke, 45*(3), 877–879.

Siegel, J. S., Seitzman, B. A., Ramsey, L. E., Ortega, M., Gordon, E. M., Dosenbach, N. U., Petersen, S. E., Shulman, G. L., & Corbetta, M. (2018). Re-emergence of modular brain networks in stroke recovery. *Cortex*, *101*, 44–59.

Simic, T., Leonard, C., Laird, L., Cupit, J., Höbler, F., & Rochon, E. (2016). A usability study of internet-based therapy for naming deficits in aphasia. *American journal of speech-language pathology, 25*(4), 642–653.

Simmons-Mackie, N., & Damico, J. S. (2001). Intervention outcomes: A clinical application of qualitative methods. *Topics in Language Disorders*, *21,* 21–36.

Skipper-Kallal, L. M., Lacey, E. H., Xing, S., & Turkeltaub, P. E. (2017). Right hemisphere remapping of naming functions depends on lesion size and location in poststroke aphasia. *Neural Plasticity*, 8740353–17.

Snell, C., Sage, K., & Lambon Ralph, M. A. (2010). How many words should we provide in anomia therapy? A meta-analysis and a case series study. *Aphasiology*, *24*(9), 1064–1094.

Stahl, B., Mohr, B., Buscher, V., Dreyer, F. R., Lucchese, G., Pulvermuller, F. (2018). Efficacy of intensive aphasia therapy in patients with chronic stroke: A randomised controlled trial. *Journal of Neurology, Neurosurgery and Psychiatry, 89*(6), 586–92.

Stockert, A., Wawrzyniak, M., Klingbeil, J., Wrede, K., Kümmerer, D., Hartwigsen, G., Kaller, C. P., Weiller, C., & Saur, D. (2020). Dynamics of language reorganization after left temporo-parietal and frontal stroke. *Brain*, *143*(3), 844–861.

Stroke Statistics. (2017). *State of the Nation, Stroke Statistics January 2017*. Retrieved from https://www.stroke.org.uk/sites/default/files

Subedi, B. S. (2004). Emerging trends of research on transfer of learning. *International Education Journal, 5*, 591-599.

Swinburn, K., Howard, D., & Porter, G. (2004). *Comprehensive Aphasia Test.* Hove: Psychology Press.

Szaflarski, J. P., Allendorfer, J. B., Banks, C., Vannest, J., & Holland, S. K. (2013). Recovered vs. not-recovered from post-stroke aphasia: The contributions from the dominant and non-dominant hemispheres. *Restorative Neurology and Neuroscience*, *31*(4), 347–360.

Szaflarski, J. P., Eaton, K., Ball, A. L., Banks, C., Vannest, J., Allendorfer, J. B., Page, S., & Holland, S. K. (2011). Poststroke aphasia recovery assessed with functional magnetic resonance imaging and a picture identification task. *Journal of Stroke and Cerebrovascular Diseases*, *20*(4), 336–345.

Szaflarski, J. P., Binder, J. R., Possing, E. T., McKiernan, K. A., Ward, B. D., & Hammeke, T. A. (2002). Language lateralization in left-handed and ambidextrous people: fMRI data. *Neurology, 39*, 238-244.

Sze, W., Hameau, S., Warren, J., & Best, W. (2021). Identifying the components of a successful spoken naming therapy: A meta-analysis of word-finding interventions for adults with aphasia. *Aphasiology, 35*(1), 33-72.

Tardif, C. L., Gauthier, C. J., Steele, C. J., Bazin, P. L., Schäfer, A., Schaefer, A…… Villringer, A. (2016). Advanced MRI techniques to improve our understanding of experience-induced neuroplasticity. *NeuroImage, 131*, 55–72.

Thomas, C., & Baker, C. I. (2013). Teaching an adult brain new tricks: A critical review of evidence for training-dependent structural plasticity in humans. *NeuroImage, 73*, 225–236.

Thomas, L., Lander, L., Cox, N., & Romani, C. (2020). Speech and language therapy for aphasia: Parameters and outcomes. *Aphasiology*, *34*(5), 603–642.

Thompson, C. K. (2017). The “right” hemisphere and language recovery in aphasia. *Paper presented at the Clinical Aphasiology Conference*. Snowbird, UT.

Thompson, C. K., & Shapiro, L. P. (2007). Complexity in treatment of syntactic deficits. *American Journal of Speech-Language Pathology, 16*(1), 30-42.

Thompson, C. (1989). Generalization research in aphasia. In T. Prescott (Ed.), *Clinical Aphasiology* (v. 18, pp. 195–222). Boston, MA: College Hill Press.

Thuillard Colombo, F., & Assal, G. (1992). Adaptation francaise du test de denomination de Boston version abrégées. *Eur. Rev. Appl. Psychol. 42*, 66–71.

Tremblay, K. L., & Kraus, N. (2002). Auditory training induces asymmetrical changes in cortical neural activity. *Journal of Speech, Language, and Hearing Research*, *45*(3), 564–572.

Turkeltaub, P. E., Messing, S., Norise, C., & Hamilton, R. H. (2011). Are networks for residual language function and recovery consistent across aphasic patients? *Neurology*, *76*(20), 1726–1734.

Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: Converging evidence from lesion and connectivity analyses. *Frontiers in System Neuroscience*, *5*, 1.

Upton, E., & Hope, T. M. H. (2020). Moving beyond the dual stream account of language. *Brain*, *143*(8), 2336–2338.

Van De Sandt-Koenderman, W. M. E., Van Harskamp, F., Duivenvoorden, H. J., Remerie, S. C.,Van Der Voort-Klees, Y. A., Wielaert, S. M.. . . . Visch-Brink, E. G. (2008). MAAS (multi-axial aphasia system): Realistic goal setting in aphasia rehabilitation. *International Journal of Rehabilitation Research, 31*(4), 314-320.

Van Hees, S., Angwin, A., McMahon, K., & Copland, D. (2013). A comparison of semantic feature analysis and phonological components analysis for the treatment of naming impairments in aphasia. *Neuropsychological Rehabilitation, 23*, 102-132.

van Hees, S., McMahon, K., Angwin, A., de Zubicaray, G., & Copland, D. A. (2014). Neural activity associated with semantic versus phonological anomia treatments in aphasia. *Brain and Language*, *129*(1), 47–57.

van Heuven, W., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67*(6), 1176-1190.

van Oers, C. A., Vink, M., van Zandvoort, M. J. E., van der Worp, H. B., de Haan, E. H. F., Kappelle, L. J., Ramsey, N. F., & Dijkhuizen, R. M. (2010). Contribution of the left and right inferior frontal gyrus in recovery from aphasia. A functional MRI study in stroke patients with preserved hemodynamic responsiveness. *NeuroImage*, *49*(1), 885–893.

Vickers, C. (2010). Social networks after the onset of aphasia: The impact of aphasia group attendance. *Aphasiology, 24*(6-8), 902-913.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Jobard, G., Petit, L., Crivello, F., … Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis*. NeuroImage, 54*(1), 577–593.

Vitali, P., Abutalebi, J., Tettamanti, M., Danna, M., Ansaldo, A.-I., Perani, D., Joanette, Y., & Cappa, S. F. (2007). Training-Induced Brain Remapping in Chronic Aphasia: A Pilot Study. *Neurorehabilitation and Neural Repair*, *21*(2), 152–160.

Wafa, H. A., Wolfe, C., Emmett, E., Roth, G. A., Johnson, C. O., & Wang, Y. (2020). Burden of stroke in Europe: Thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years. *Stroke, 51*(8), 2418-2427.

Walker, G. M., Schwartz, M. F., Kimberg, D. Y., Faseyitan, O., Brecher, A., Dell, G. S., & Coslett, H. B. (2011). Support for anterior temporal involvement in semantic error production in aphasia: New evidence from VLSM. *Brain and Language*, *117*(3), 110–122.

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25***,** 1256– 1271.

Wambaugh, J., Linebaugh, C., Doyle, P., Martinez, Z., Kalinyak-Fliszar, M., & Spencer, K. (2001). Effects of two cueing treatments on lexical retrieval in aphasic speakers with different levels of deficit. *Aphasiology*, *15*, 933–950.

Wan, C. Y., Zheng, X., Marchina, S., Norton, A., & Schlaug, G. (2014). Intensive therapy induces contralateral white matter changes in chronic stroke patients with Broca's aphasia. *Brain and Language, 136*, 1–7.

Watila, M. M., & Balarabe, S. A. (2015). Factors predicting post-stroke aphasia recovery. *Journal of the Neurological Sciences, 352*(1–2), 12–18.

Webster, J., Whitworth, A., & Morris, J. (2015). Is it time to stop "fishing"? A review of generalisation following aphasia intervention. *Aphasiology, 29*(11), 1240-1264.

Wernicke, C. (1874). *Der aphasiche Symptomenkomplex*. Breslau, Poland: Cohen and Weigert.

Whitworth, A., Webster, J., & Howard, D. (2014). *A Cognitive Neuropsychological Approach to Assessment and Intervention in Aphasia*. London: Psychology Press.

Wilke, M., de Haan, B., Juenger, H., & Karnath, H. O. (2011). Manual, semi-automated, and automated delineation of chronic brain lesions: a comparison of methods. *Neuroimage, 56*(4), 2038-2046.

Wilson, S. M., & Hula, W. D. (2019). Multivariate Approaches to Understanding Aphasia and its Neural Substrates. *Current Neurology and Neuroscience Reports*, *19*(8), 53–59.

Winhuisen, L., Thiel, A., Schumacher, B., Kessler, J., Rudolf, J., Haupt, W. F., & Heiss, W. D. (2005). Role of the contralateral inferior frontal gyrus in recovery of language

function in poststroke aphasia: A combined repetitive transcranial magnetic stimulation and positron emission tomography study. *Stroke (1970)*, *36*(8), 1759–1763.

Wise, R. (2003). Language systems in normal and aphasic human subjects: functional imaging studies and inferences from animal studies. *British Medical Bulletin*, *65*(1), 95–119.

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology, 23*, 1338-1352

Woodhead, Z., Crinion, J., Teki, S., Penny, W., Price, C. J., & Leff, A. P. (2017). Auditory training changes temporal lobe connectivity in 'Wernicke's aphasia': A randomised trial. *Journal of Neurology, Neurosurgery & Psychiatry, 88*(7), 586-594.

Woolf, C., Caute, A., Haigh, Z., Galliers, J., Wilson, S., Kessie, A., et al. (2016). A comparison of remote therapy, face to face therapy and an attention control intervention for people with aphasia: a quasi-randomised controlled feasibility study. *Clin. Rehabil.* 30, 395–373.

Worrall, L., Hudson, K., Khan, A., Ryan, B., & Simmons-Mackie, N. (2016). Determinants of living well with aphasia in the first year post-stroke: A prospective cohort study. *Archives of Physical Medicine and Rehabilitation*, *98*(2), 235–240.

Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25,* 309–322.

Xing, S., Lacey, E. H., Skipper-Kallal, L. M., Jiang, X., Harris-Love, M. L., Zeng, J., & Turkeltaub, P. E. (2016). Right hemisphere grey matter structure and language outcomes in chronic left hemisphere stroke. *Brain, 139*(1), 227–241.

Xue, G., Chen, C., Jin, Z., & Dong, Q. (2006). Language experience shapes fusiform activation when processing a logographic artificial language: An fMRI training study. *Neuroimage, 31*(3), 1315-1326.

Yang, H. (2013). The case for being automatic: Introducing the Automatic Linear Modeling (LINEAR) procedure in SPSS statistics. *Multiple Linear Regression Viewpoints, 39*(2), 27 - 37.

Yeung, O., & Law, S. P. (2010). Executive functions and aphasia treatment outcomes: Data from an ortho-phonological cueing therapy for anomia in Chinese. *International Journal of Speech-Language Pathology, 12*, 529-544.

Yourganov, G., Smith, K. G., Fridriksson, J., & Rorden, C. (2015). Predicting aphasia type from brain damage measured with structural MRI. *Cortex, 73*, 203–15.

Zhang, J., Zhang, Y., Xing, S., Liang, Z., & Zeng, J. (2012). Secondary neurodegeneration in remote regions after focal cerebral infarction: a new target for stroke management? *Stroke, 43*, 1700–1705.

Zheng, C., Lynch, L., & Taylor, N. (2016). Effect of computer therapy in aphasia: A systematic review. *Aphasiology, 30*(2-3), 1-34.

8 Appendices

8.1 Appendix A: Speech error coding

Response Category	Description	Examples of Target and Responses
Correct	Correct production of the target word (including production of the target word followed by an incorrect response).	*Fish*: "fish" hmm..fish"; "fish…whale'
Self-correction	Production of an incorrect word followed by the correct production of the target word.	*Fish*: "cat no no.. fish"
Late	Correct production of the target word but word produced in the recording for the following picture.	*Fish:* "oh it's erm.." *Football:* "fish"
Morphological error	Production of the target word but with an incorrect morphological inflection or omission of the correct morphological inflection.	*Run*: "running"; "runs" *Sitting*: "sit"
Non-response	Filler words or silence.	"No.. I don't know"; "mmm"
Synonym	A culturally acceptable response to the picture.	*Little*: "small" *Talking*: "speaking"
Semantic	A coordinate, superordinate or subordinate of the target word, or response highly associated to the target word.	*Fish:* "cod" *Shepherd*: "crook"
Unrelated	A real word response that does not have a semantic or phonological relation to the target word.	*Fish*: "radio"
Phonemic	A real word or non-real word response that preserves at least 50% of the segments from the target word. This includes the substitution, omission or addition of phonemes.	*Fish*: "bish"; "fi"; "fishk"
Neologism	A non-real word that does not fulfil the above criteria for a phonemic error.	*Fish*: "zizi"
Perseveration	An unrelated response (real word or non-real word) that has been repeated three or more times previously.	*Fish:* "car" *Football:* "car" *Boots:* "car" *Table:* "car"
Visual	A response that looks visibly similar to the target but is not semantically related. Also, for words that are abstract, this includes responses that describe the picture but are not related to the target word.	*Football*: "globe" *Important*: 'letter'
Circumlocution	A multi-word response that is an informative description of the picture, without producing the target word.	*Fish*: "the thing that swims in the sea'
Mixed	A response that is a mixture of either visual, semantic or phonological errors.	*Boots*: "choes" *Football*: "klobe"
Partial phonemic	A response that has at least one phoneme that is related or unrelated to the target word.	*Fish*: "/s/"

8.2 Appendix B: Spoken Picture Description and associated items from the WRT

Scene 1

Neighbour	Fork
Lawnmower	Apple
Castle	Picture
Crown	Frame
Flag	Door
Sand	Hair
Dinner	White
Carrots	Table
Juice	Alcohol
Soup	Knife
Shouting	Summer
(She)	(He)

Scene 2

Motorbike	Train
Girl	Ambulance
Wheel	Nurse
Snow	Coat
Police	Ring
Student	Clock
Water	Street
Glasses	Shops
Pizza	Cold
Talking	Writing
Sitting	Winter
(She)	(He)