TOWARDS A NEW COGNITIVE NEUROSCIENTIFIC ACCOUNT OF VERBAL SERIAL SHORT-TERM MEMORY AND LEARNING: A PERCEPTUAL-MOTOR APPROACH

Jasmine K. Virhia

Thesis submitted in fulfilment of the

degree of Doctor of Philosophy

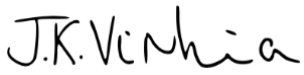
Department of Psychology

Royal Holloway, University of London

June 2021

DECLARATION OF AUTHORSHIP

I, Jasmine Kaur Virhia, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed: 

Date: 10th June 2021

ABSTRACT

Verbal serial short-term memory and learning have, classically, been attributed to the action of a specialised phonological short-term store localised to the inferior parietal lobe. Recently, however, the behavioural hallmarks of a phonological store have been reinterpreted in terms of articulatory planning and perceptual organisation processes. Two fMRI experiments therefore examined whether patterns of neural activation during verbal serial recall can be explained in terms of articulatory planning processes (instantiated via cortico-cerebellar loops) and modality-specific perceptual processes, without invoking a non-motoric and modality-independent phonological store. It was also examined whether the long-term learning of a verbal sequence is associated with a down-regulation of activity in the cerebellum, consistent with research on motor skill automatisation. Activations associated with articulatory rehearsal of a sequence of letter-names during a retention interval were observed in prefrontal area 9/46, the premotor cortex, and cerebellar lobules HVI and HVIIB—all previously implicated in motor planning—with both auditory (Experiment 1) and visual presentation (Experiment 2). Additionally, activation tied specifically to auditory presentation was observed in the premotor cortex, primary auditory cortex and planum temporale, while activation in the premotor cortex, temporo-occipital-fusiform cortex, intraparietal sulcus, Crus II and lobule HVIIB was observed during visual presentation. Contrary to the phonological store theory, no region in the inferior parietal lobe was found to be active during both stimulus-presentation and rehearsal. Instead, different subdivisions of the intraparietal sulcus were associated with the rehearsal of auditory sequences and with the presentation and rehearsal of visual sequences. Only motor-planning regions were consistently active across both input-modalities and trial-phases, whilst distinct regions were active during presentation as a function of input-modality. While

behavioural evidence of sequence learning was observed with both auditory (Experiment 1B) and visual presentation (Experiment 2B), changes in neural activity appeared to reflect general task-set, rather than sequence-specific, learning.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisors, Dr Rob Hughes and Professor Narender Ramnani. Having started my academic journey as an undergraduate studying English Literature and Linguistics, never would I have imagined completing a PhD in Cognitive Neuroscience. Thank you for going on this journey with me. I'm honoured to have conducted the *first* neuroimaging investigation within the perceptual-motor framework, I hope the work in this thesis forms the basis of many more investigations to come! I'd also like to thank Dr Tibor Auer and Dr Craig Arnold for their expertise in programming. They both provided me with a true admiration for the magic that is code and enabled me to design and analyse this huge project. To Ari Lingeswaran, thank you for all your assistance in the MRI unit-without those participants there would be no PhD!

I'm incredibly grateful to have shared this experience with so many other PhD students and to have formed some incredible friendships during my time at Royal Holloway: A special shout out to Adam, Kathrin, Farah, Lena, Sahil and the Perception, Action & Memory Lab. In particular, I'd like to thank Dr Jennifer Mills for all of her emotional and scientific support over the past few years. I wouldn't have survived this PhD if it weren't for Jen. We started off as colleagues, sharing very similar experiences in our academic journeys and overlapping in a lot of our research. Now however, I'm lucky enough to say she one of my closest and most treasured friends. We owe so much to Sarah Bellum for bringing us together. After almost a decade of friendship I'd like to thank Simran, simply for being my best friend and for everything that role entails- the good, the bad and the ugly. Thank you for always having my back. I wouldn't even know where to begin but I am so very grateful for

her and her excellent choices in fancy restaurants. Here's to travelling the world once again (COVID-19 pending).

To my parents, Perminder and Leonardo, my Nanaji and Naniji (Santokh Singh & Kulwant Kaur), and uncle Mandeep, I owe everything. This PhD is a testament to the support of my family. The past eighteen months especially, have required me to exercise incredible patience and resilience- characteristics my Nanaji so effortlessly embodies. I hope reaching the height of my education has made his sacrifices early on in life worth it. Those who have the pleasure of knowing my Mum know that she exudes kindness, positivity and strength, especially when faced with adversity. Words fail me when I try to express my gratitude or explain how she constantly inspires me. All I can say is thank you for being you, Mum. Your unwavering belief in me and constant reminders of my capabilities have kept me going through some very difficult times. I am beyond blessed to be your daughter and I hope to always make you proud.

I am a fundamentally different person to who I was when I began this PhD. I have learned just as much about myself and life as I have science over the past four years and will carry this experience with me for a very long time to come.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: GENERAL INTRODUCTION

1.1 Introduction

The retention and reproduction of verbal sequences in the short term (in the order of a few seconds) is imperative for comprehensive linguistic communication (reading, writing, speaking, signing, seeing and listening; e.g., Baddeley, 1989) and the ability to retain novel sequence elements during linguistic perception and production over the long-term is a key aspect of language acquisition (Baddeley et al., 1998; Page & Norris, 2009a,b; Saffran et al., 1997). In particular, the process of learning novel verbal sequences—where short-term sequence representations are converted into long-term ones is a critical process in the learning of new word-forms (e.g., Burgess & Hitch, 2005; Szmalec et al., 2009).

The dominant cognitive account of verbal serial short-term memory and sequence learning is, arguably, the phonological loop model (Baddeley, 1986, 2007). The phonological loop is sub-divided into two components: the *phonological store,* a discrete cognitive module used to store verbal information for around 2 s and an articulatory control process used to refresh decay-prone items in the store and to convert visual-verbal material into phonological form. Although supported by an articulatory process, the store is assumed to be both functionally and anatomically separate from articulatory, perceptual, and language processing systems (Baddeley, 1988, 1989, 2012; Baddeley et al., 1984). Indeed, of particular interest in the present thesis is that the phonological loop model is perhaps the most famous embodiment of the more general cognitivist view that holds that motoric and perceptual processes are peripheral to the supposedly 'central' cognitive structures that subserve fundamental psychological functions such as memory, language and learning (see, e.g., Hurley,

2001). In the phonological loop model, then, motoric and perceptual processes merely serve as a means of outputting information from, and providing the input to, a phonological store (Baddeley, 1986, 2012).

The impetus for the present research is recent evidence suggesting that the supposed empirical signatures of a phonological store can be explained more parsimoniously by recourse to articulatory planning and perceptual organisation processes without having to invoke a dedicated memory module and that, more generally, motor and perceptual processes play an instrumental, not peripheral, role in verbal serial short-term memory and learning (Jones et al., 2006, 2004; Jones & Macken, 2018; Sjöblom & Hughes, 2020). The approach taken in the present thesis was to re-examine the cognitive neuroscience of verbal short-term memory, which has tended, like cognitive psychology, to be dominated by the classical phonological store theory (Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996). Indeed, the cognitive neuroscientific evidence has typically been interpreted as reinforcing the concept of a functionally distinct phonological store on the grounds that a phonological store can, it has been argued, be localised to an *anatomically* distinct region of the brain, separate from brain regions involved in articulatory, perceptual, or general language functions (e.g., Baddeley, 2007, Buchsbaum & D'Esposito, 2008). Motivated by the reconceptualisation of the cognitive psychology of verbal serial short-term memory performance in motor and perceptual terms (e.g., Jones et al., 2004), the current research uses functional magnetic resonance imaging (fMRI) to examine the extent to which the neural bases of such performance—as well verbal sequence learning—can instead be understood primarily in terms of the action of brain regions and systems that support articulatory planning, perceptual and perceptual-motor mapping processes, and motor skill learning.

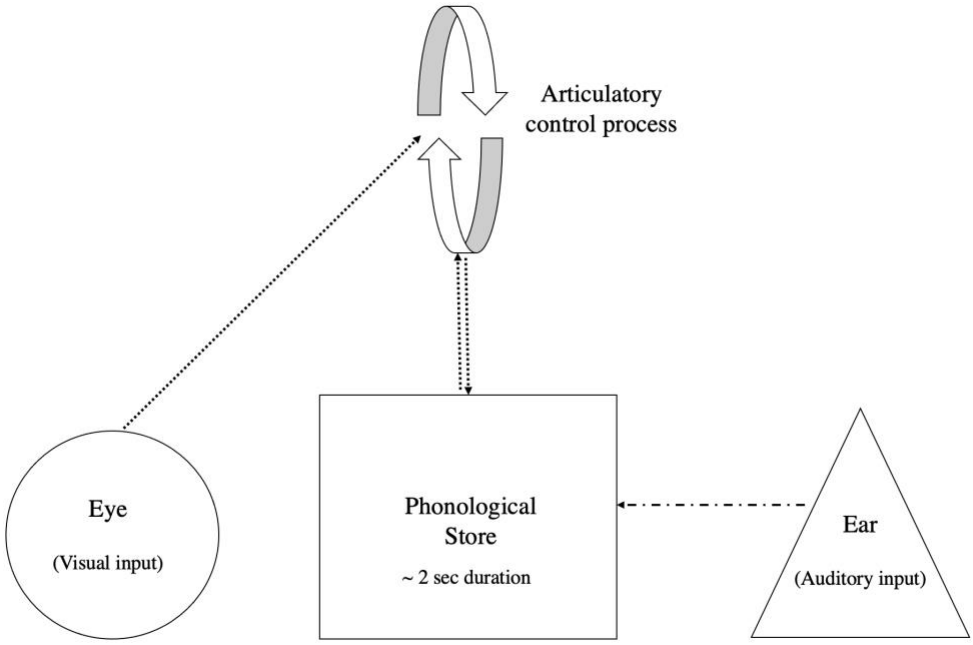
1.2 The Phonological Loop Model

The phonological loop model is part of the working memory model which grew out of the modal model of memory (Atkinson and Shiffrin, 1968; Norman, 1970). The modal model refers to a set of theoretical assumptions shared by a number of early accounts of the overall architecture of the memory system in which there were distinct (though interacting) short-term and long-term memory structures (Baddeley et al., 2019; Norman, 1970). The modal model also distinguished between high-capacity pre-categorical sensory registers with fast decay rates and a unitary short-term memory store with a severely limited capacity (cf. Miller, 1956). The store was served by the pre-categorical sensory registers and rehearsal of the contents of the short-term store led to long-term memory formation. From the concept of a short-term memory store within the modal model arose the working memory model (Baddeley & Hitch, 1974). Here, the short-term store component within the modal model was expanded into a 'working memory' system—comprising active control functions as well as passive temporary storage—that supports the performance of complex tasks such as comprehension and learning (Baddeley, 1983, 2000; Baddeley & Hitch, 1974). The phonological loop is separate from the visuo-spatial sketchpad, another short-term memory module within the working memory model dedicated to the storage of visuo-spatial input. The central executive acts as a control system for organisational functions, directing attention to resources for input into the phonological loop or visuo-spatial sketchpad (Baddeley, 2012). Another limited capacity module controlled by the central executive is the episodic buffer. The buffer is capable of utilising a multimodal code and acts as an interface between the short-term stores and long-term memory (Baddeley, 2000).

The phonological loop construct was developed primarily on the basis of performance in short-term verbal serial recall, a task where participants are presented with a sequence of around 5-9 verbal items (e.g., letters, words, or digits) at a rate of around one or two per second. Following sequence presentation, or after a short temporal delay following the last item, participants are asked to recall the sequence in strict serial order either in written, spoken or typed form, or by clicking the (re-presented) items in the correct order. As shown in Figure 1, the phonological loop model posits that the items enter a passive phonological store for a duration of around 2 s before they decay. The items can be refreshed, however, by an articulatory control process that serves to prevent the loss of items from the store (articulatory rehearsal). Importantly, auditory-verbal (i.e., spoken) input and visual-verbal input gain access to the phonological store via different routes: Auditory-verbal information has direct and obligatory access to the store because it is deemed to be in a code already suitable for the phonological store. Visual-verbal information, however, must be converted into the store's code indirectly via a grapheme-phoneme conversion process, another function of the articulatory control process (Baddeley, 2003). Note that it is the fact that auditory-verbal and visual-verbal inputs access the same store (albeit via different routes) that makes it a *phonological* store: the representations therein are modality-independent representations devoid of sensory or articulatory features.

Figure 1

A Schematic Diagram of The Phonological Loop Model (e.g., Baddeley, 1986, 2007).



The fractionation of the phonological loop into a passive phonological store and an articulatory control process was based in part on neuropsychological evidence that appeared to show that verbal short-term memory deficits could be observed in the absence of speech production deficits. For example, patient J. B. (Warrington and Shallice, 1969) was able to perceive, comprehend and produce language without the kind of abnormal pauses or errors that would indicate aphasia. It was thus inferred that language faculties necessary for "online" comprehension and production of meaningful speech are functionally dissociable from verbal short-term memory (Shallice & Butterworth, 1977). More specifically, it was argued that these patients suffered from a deficit to an auditory-verbal short-term store (Warrington & Shallice, 1969) that temporarily stores information that arrives therein as the output of a separate speech perception system. It was suggested that the memory impairment could not be

explained by a general auditory or general memory deficit as J. B. was able to complete short-term memory tasks involving visuospatial information or non-linguistic auditory information (Shallice & Butterworth, 1977).

The putative empirical hallmark of the passive phonological store component is the phonological similarity effect: Short-term serial recall of phonologically similar items (e.g., letters such as 'P, V, T, C,' or the words *man, mat, can, map*, and *cat*) is poorer than for dissimilar items such as 'F, K, Q, H' (or *pit, day, cow, pen,* and *sup*; Baddeley, 1966; Coltheart, 1993; Conrad & Hull, 1964). This effect is found with both auditory and visual presentation of the items and is thought, on the phonological loop model, to be caused by confusions between partially overlapping representations during retrieval of phonologically similar items from the phonological store, leading to serial order errors (Baddeley, 1989; Baddeley & Larsen, 2007). Interestingly, given the framework to be adopted in the present thesis, proponents of the phonological loop model suggested initially that the phonological similarity effect reflects the action of the articulatory system. Consistent with this, if participants are required to vocalise (or whisper) an irrelevant word or sequence (e.g., "the, the, the…")—so-called *articulatory suppression*—so that articulatory rehearsal of those items is restricted, the phonological similarity effect is removed (Baddeley et al., 1984; Murray, 1968). As a result, the phonological loop was originally termed the *articulatory loop* (Baddeley et al., 1984). However, apparently pointing away from an articulatory account, it was then observed that the elimination of the phonological similarity effect under suppression only occurs with visual presentation of the items; the phonological similarity effect survives articulatory suppression if the items are presented auditorily (Baddeley et al., 1984; see also Murray, 1968). Thus, what was originally termed the articulatory loop "was renamed the phonological loop, on the grounds that the capacity

for storage was the central feature of the system, which can operate without articulation, provided material is presented auditorily" (Baddeley and Larsen, 2007, p. 497). On the updated phonological loop model, then, the phonological similarity effect disappears under articulatory suppression with visual presentation because articulatory processes are required to give visual-verbal items (but not auditory-verbal items) access to the phonological store, not because of their involvement in rehearsal (e.g., Baddeley, 2007; cf. present Figure 1). It is the survival of the phonological similarity effect under suppression with auditory input that serves as arguably the strongest evidence for the notion of a passive phonological store—to which auditory (but not visual) input gains obligatory access—separate from articulatory processes (Baddeley et al., 1984).

1.3 Verbal Sequence Learning

While the phonological store construct was developed originally on the basis of the study of short-term serial recall and seemed to account well for a range of serial recall phenomena, it soon became unclear what the evolutionary function of the loop was: Patients with apparently pure deficits of phonological storage (Shallice & Butterworth, 1977; Warrington and Shallice, 1969) suffered few problems navigating their everyday lives (Baddeley et al., 1988) but did, however, have difficulty with understanding long and complex sentences (Vallar & Baddeley, 1987). This initially suggested that the store might exist as a fulcrum for comprehension of complex sentences, but this did not appear to be an important enough function to have led to the evolution of a dedicated phonological store. In continuing the search for an evolutionary function for the store, the idea emerged that it was to support key aspects of language acquisition. Specifically, it has been argued that the evolved function of the phonological store is to support the long-term learning of a novel sequence of

linguistic elements, that is, word-form learning (Baddeley et al., 1988). That is, the phonological store temporarily holds the already-known individual phonemes making up a newly encountered word while permanent memory traces of their novel order are formed (Atkins & Baddeley, 1998; Baddeley et al., 1998; Gathercole, 2006). Thus, the use of temporary phonological storage for short-term recall is now considered a by-product rather than the evolved purpose of the store (Atkins & Baddeley, 1998; Baddeley, 2012; Baddeley et al., 1998).

In recent decades, the most prominent way of testing the view that the phonological store supports language learning has been through the study of the Hebb effect (Hebb, 1961). This refers to the improved short-term serial recall of a sequence—most typically a sequence of verbal items such as letter-names or digits—that is repeated every few trials (e.g., every third sequence) compared to non-repeating 'filler' sequences. The increase in serial recall accuracy is observed even if subjects are unaware of the repetition (McKelvie, 1987; Stadler, 1993). The effect is therefore taken to indicate the incidental long-term learning of the repeating sequence (e.g., Hebb, 1961; Melton, 1963; Mosse & Jarrold, 2008; Hitch et al., 2009; Page & Norris, 2009a; Sjöblom & Hughes, 2020; St-Louis et al., 2019; Szmalec et al., 2012; Szmalec et al., 2011; Yanaoka et al., 2019).

A number of findings suggest that Hebb sequence learning in verbal serial recall is a valid analogue of word-form learning (Page & Norris, 2008; Norris et al., 2018; Szmalec et al., 2009). Indeed, it seems reasonable to assume that the mechanisms relied upon for learning a novel sequence of letters or words presented for short-term serial recall would not be entirely distinct from those necessary for learning a sequence of phonemes or syllables that form a newly encountered word (e.g., learning the list of letters "B, J, F, M, L" could be seen as being akin to learning

the new word “Beejayeffemelle”; Page & Norris, 2009b). This interpretation is further supported by the findings that verbal serial recall performance shows similar empirical patterns to immediate nonword repetition (Gupta, 2005) and that a Hebb effect is found when the repeating sequence is a nonword (Hughes et al., 2021).

Phonological-store based computational models account for both the short-term ordering of a verbal sequence (e.g., for serial recall) and its long-term learning (cf. the Hebb effect) by positing an abstract (non-motoric) sequencing mechanism that acts upon item-representations held in the phonological store (Burgess & Hitch, 2005; Page & Norris, 2009b). Of most relevance to the present thesis is that articulatory processes in these models, as in the phonological loop model upon which they are based, serve only to reactivate and refresh abstract phonological representations of individual items, or to convert visual input into phonological form. That is, the articulatory processes are described as having no direct role in serial ordering (again as in the original phonological loop model), rather, there is a separate mechanism responsible for serial order. In this view, serial order is represented either in terms of a primacy gradient of item-activation strengths (the item still most active at retrieval will be identifiable as having occurred first in the sequence, the second-most active as having occurred second, and so on; Page & Norris, 2009b) or via item-(absolute)position connection-weights (Burgess & Hitch, 1999, 2006). Indeed, the models have a two-stage architecture such that representations of the phonology of verbal input held within the passive phonological store (and supported by articulatory processes) is independent of a second stage at which those representations are linked to the abstract mechanism that represents their serial order (Burgess & Hitch, 2005; Page & Norris, 2009b). Thus, whilst the phonological store is critical for temporarily holding representations of individual verbal items, and articulatory rehearsal supports

this function, it is the abstract ordering mechanism that drives short-term verbal serial recall and long-term sequence learning. The structural separation of item-identity and the ordering mechanism in these models predicts, therefore, that variables assumed to affect short-term serial recall by impairing phonological short-term storage as opposed to the ordering mechanism—such as phonological similarity and articulatory suppression—should not affect long-term sequence learning. In line with this prediction, both Hitch et al. (2009) and Page et al. (2006) reported that while articulatory suppression had the usual marked effect on short-term serial recall, it did not modulate Hebb sequence learning. Hitch et al. (2009) showed in addition that phonological similarity also had its usual detrimental effect on serial recall but no effect on Hebb sequence learning.

1.4 Challenges to the Concept of a Phonological Store and Development of an Alternative, Perceptual-Motor, Approach to Verbal Serial Short-Term Memory and Learning

Challenges to the concept of a phonological store over the past fifteen years or so have pointed towards the centrality of perceptual and motor processes in verbal serial short-term memory and learning. As described earlier (Section 1.2), crucial to the notion of a passive phonological store separate from articulatory processes is the finding that the main empirical signature of the store—the phonological similarity effect—is still found even when articulatory processes are impeded via articulatory suppression so long as the items are presented auditorily and hence enjoy direct access to the store (Baddeley et al., 1984). On closer examination, however, this interaction is not as originally assumed: It transpires that the survival of the phonological similarity effect under suppression with auditory lists is restricted to recency (the last one or two items in the list). That is, the phonological similarity effect is indeed eliminated by

articulatory suppression throughout most of the serial position curve even with auditory presentation. Furthermore, the survival of the effect at recency under suppression with auditory input has been shown to be a consequence of the *modality effect,* the enhanced recall of the last one or two items of an auditorily presented compared to a visually presented sequence (Conrad & Hull, 1968). It is the fact that the modality effect is only found (or is at least much more pronounced) with phonologically dissimilar compared to similar sequences (Crowder, 1971; Jones et al., 2004; Watkins et al., 1974) that accounts for the survival of the phonological similarity effect under suppression with auditory presentation. This residual effect can be accounted for in terms of the action of passive acoustic-based perceptual organisation but not phonological storage (Jones et al., 2004; Nicholls & Jones, 2002). Thus, when stimuli are presented auditorily and articulatory planning is impeded, the residual 'phonological similarity effect' has been shown to be an *acoustic similarity effect* (e.g., Jones et al., 2006, 2004). Indeed, on the Working Memory model, the modality effect is considered "peripheral to the working memory system" (Baddeley, 1986, p. 95) and hence unrelated to the phonological store. Thus, the three-way interaction that was "crucial to separating the two components of the articulatory loop, the phonological store and the articulatory control process" (Baddeley, 1986, p. 257) does not take the form that was used to justify such a separation.

The fact that, aside from the residual acoustic-based effect at recency, engagement in articulatory processing is, after all, a prerequisite for the phonological similarity effect locates that effect in the articulatory system, not a passive phonological store. More specifically, it has been suggested that when subjects are free to engage in articulatory rehearsal, the phonological similarity effect is the result of sub-lexical speech planning errors (Acheson & MacDonald, 2009; Ellis, 1980; Page

et al., 2007) and should therefore be understood as an *articulatory similarity effect.* For example, the errors made when recalling a phonologically similar list are systematic and strikingly similar to those occasionally produced during normal phrase or sentence production (cf. the *error equivalence hypothesis*; Ellis, 1980). For example, Ellis (1980) presented five syllables auditorily with different combinations of vowels and consonants for serial recall and observed errors that were analogous to errors found in naturally produced language. For instance, the errors were dependent on contextual similarity where two consonants are more likely to swap places if the syllables of which they are a part share a similar or identical vowel, e.g., when intending to say "light a fire", 'l' and 'f' are swapped and the utterance becomes "fight a liar" (MacKay, 1970; Nooteboom, 1967). This is also an example of the effect of within-syllable position, where phonemes exchange but retain their respective positions (the onsets of each word in the foregoing example) in the syllables to which they have migrated (Boomer & Laver, 1968; MacKay, 1970; Nooteboom, 1967; Levelt, 1989). The typical phonologically similar list in a serial recall tasks (e.g., B, C, D, G, P, T, V) exhibits maximal contextual similarity: All the items have different consonant onsets but share the same vowel (/i/ "ee") which is also the coda in every item. The fact that manual gestures also show errors due to 'phonological' similarity (Leybaert & Lechat, 2001; Wilson & Fox, 2007) further supports the notion that the vocal phonological similarity effect is rooted in the articulatory gestures of speech and has a motoric basis operating on phonetic not phonological elements.

The reinterpretation of the phonological similarity effect in articulatory and acoustic-perceptual terms has played a key role in the development of a perceptual-motor account of verbal serial short-term memory more generally. This account proffers that performance in the verbal short-term memory setting, irrespective of

presentation modality (visual or auditory), is parasitic on the formation of an articulatory plan for task-relevant motor outputs and processes involved in the perceptual organisation of auditory input (when items are auditorily presented; Hughes et al., 2009, 2016; Hughes & Marsh, 2017; Jones et al., 2006, 2004; Macken et al., 2016; Sjöblom & Hughes, 2020). This view holds that the short-term retention of a sequence is supported primarily by the motor-plan itself and denies the need to invoke a dedicated phonological store, thereby offering a more parsimonious way of understanding short-term memory performance (e.g., Jones et al., 2007). On this account, successful recall and learning are products of object-oriented processes that organise the environment into perceptual objects facilitating control programs for goal-directed actions. This approach maintains that verbal material should be viewed in the same way as other kinds of material. The way in which people therefore encounter and manipulate information, including verbal information, should be understood with respect to the processes that organise the material into perceptual objects that may then be apprehended and manipulated by bodily effector systems (Jones & Macken, 2018).

A key distinction between the perceptual-motor account and the phonological loop model is the function ascribed to articulatory processes. In the phonological loop model, their function is primarily to offset another, negative, process: to refresh item representations in the face of decay within a discrete verbal store (as well as to convert visuo-verbal items into a phonological form; Baddeley, 2007). In the perceptual-motor account, in contrast, articulatory rehearsal or *planning* is (re)conceptualised as playing the constructive role of binding otherwise sequentially unrelated items together to form a fluid vocal-motor object. To elaborate, the key phenomena upon which the majority of short-term memory theorising are based, are observed when there is little—if any—

burden on recalling the items themselves. For example, a closed set of items is often used (e.g., digits 1-8) where it is only their order that changes across trials. The order of items is not constrained by any form of semantic category, syntactic or grammatic rules, nor does the to-be-recalled sequence hold any supra-item meaning.

Additionally, in some studies, the items themselves are re-presented to subjects at the recall stage and hence, again, the key requirement is to reproduce them in the correct order (*order reconstruction*). It is the fact that items are presented in an unfamiliar order that demands some sort of active short-term processing such as the assembly of a motor plan. As the items in a serial recall task are, by design, unrelated, they must somehow be made to relate to one another 'on-the-fly'. The articulatory planning system is therefore co-opted on account of the fact that the inherently sequential and temporally structured nature of an articulatory plan affords a ready means of creating a singular "motor object" from the otherwise discrete, unrelated, elements that make up the list. For example, speech habits such as the imposition of prosody and co-articulation imbue the material with new sequential information that is not present in the list itself (e.g., Woodward et al., 2008). In this view, then, short-term retention and reproduction of a sequence is a by-product of opportunistically hijacking a process that has evolved for the production of fluent articulatory gestures, not the result of the action of a bespoke memory system. This view directly contrasts with the role ascribed to articulatory processes in the phonological loop model where articulation supports item-encoding (in the case of visual-verbal input) and item-storage but not item-order processing, either for short-term recall or for long-term learning (Baddeley, 2007; Burgess & Hitch, 2005; Hitch et al., 2009; Page et al., 2006,2009a,b).

The different function attributed to articulatory processes in the perceptual-motor account compared to the phonological loop model has important implications for understanding the role of presentation-modality (auditory and visual) in verbal serial short-term memory. In the phonological loop model, modality only influences serial short-term memory insofar as auditory input gains obligatory access to the phonological store whilst visual input must undergo a grapheme-phoneme conversion process. Aside from this, sensory perceptual processes—and hence the different forms these take depending on input-modality—more generally are not important in this model as the core of the model is the *phonological* store which, as already described, is separate from perception (as well as from motor processes). In the perceptual-motor account, in contrast, the particular input-modality has important ramifications for the ease with which perceptual input can be mapped onto an articulatory plan (e.g., Hughes et al., 2009, 2016; Macken et al., 2016). This is best illustrated perhaps through the *inverted modality effect* (Macken et al., 2016). This refers to the observation that the advantage for auditory sequences at recency (the classic modality effect; Crowder, 1971) coincides with a serial recall advantage with visual lists at pre-recency (mainly at mid-list items when performance has come away from the ceiling; Beaman, 2002; Grenfell-Essam et al., 2017; Macken et al., 2016). While auditory recency is readily attributable to acoustic-based perceptual organisation processes associated with the salience of perceptual-boundary information (Bregman & Rudnicky, 1975; Nicholls & Jones, 2002), the inverted modality effect—which, unlike auditory recency, is abolished under articulatory suppression—can be attributed to the greater difficulty of mapping an auditory-verbal compared to a visual-verbal sequence onto an articulatory plan (Macken et al., 2016). The perceptual-motor mapping account of the inverted

modality effect begins with the observation that there are stark differences in the way perceptual organisation operates in the auditory and visual modalities.

A sequence of auditory-verbal items (presented in a common voice and from a common location as in almost all serial recall studies; cf. Hughes et al., 2009, 2016) will tend to be pre-attentively grouped together to form a temporally-extended perceptual object (or 'stream'; Bregman, 1990). While this object—by virtue of being such—will have strong boundaries (hence accounting for the classic modality effect), the elements within the object become less perceptually salient. In contrast, visual information is perceptually organised on a different basis: The visual perceptual system is not geared for streaming in time, rather it is mechanised to group information across space (Bregman, 1990). In comparison to auditory items comprising a stream (presented in a single location), visual items within a list are relatively perceptually independent, thus making perceptual boundaries of the list less salient (Bregman & Rudinicky, 1975). Instead, individual visual items can be attended to more specifically, rather than the list as an entire object and the inverted modality effect reflects the propensity of visually presented (and temporally unbound) items to integrate into a motor-plan (Macken et al., 2016; see also Grenfell-Essam et al., 2017).

The notion that there is no phonological store begs the question of what, therefore, supports the function for which the phonological store is said to have evolved, namely, verbal sequence learning (Baddeley et al., 1998)? Within the perceptual-motor framework, it has begun to be argued that verbal sequence learning may be largely a by-product of the articulatory planning processes engaged during attempts to retain and reproduce a sequence over the short term. In the context of verbal sequence learning, the repetition of an articulatory motor-plan reinforces co-articulation, i.e., the adaptation of how verbal items are uttered (or planned to be

uttered) depending on what is to be uttered next (Sternberg et al., 1980) meaning that a vocal motor plan becomes increasingly fluent and the occurrence of item-order errors reduces. It would therefore be expected that when based on motor infrastructure (as bodily effector systems are), that the acquisition of long-term motor-skills emerges from short-term motor control (Willingham, 1998).

At first glance, the evidence discussed earlier suggesting that articulatory processes are not involved in Hebb verbal sequence learning would appear to already refute a motor-based account of such learning (Hitch et al., 2009; Page et al., 2006). However, this evidence has recently come under scrutiny (Sjöblom & Hughes, 2020). To re-cap, both Hitch et al. (2009) and Page et al. (2006) found that articulatory suppression affected short-term serial recall but did not modulate Hebb sequence learning. Hitch et al. (2009) also reported that while phonological similarity had its usual detrimental effect on serial recall, it did not affect Hebb sequence learning. However, the results of Sjöblom and Hughes (2020) suggest that these conclusions were premature and that articulatory planning does indeed play a key role not only in short-term verbal serial short-term memory but also long-term verbal sequence learning: They found that Hebb sequence learning of a visually-presented sequence was indeed markedly attenuated when articulatory planning of the Hebb sequence was restricted by articulatory suppression. This was also the case with auditory sequences but less so, suggesting that passive perceptual organisation processes can support learning of auditory sequences to some extent when articulatory planning is impeded. It was also found that phonological similarity modulates Hebb sequence learning: Sequences for which articulatory planning would be expected to be relatively disfluent (phonologically similar sequences) gained more from Hebb repetition than sequences whose planning would be expected to be relatively fluent from the outset

(phonologically dissimilar sequences). Again, this differed according to input-modality: the enhanced learning effect was only found with visual sequences when passive auditory perceptual organisation processes could not contribute to the learning. Sjöblom and Hughes (2020) suggested that the discrepancy between their results and those from previous studies (Hitch et al., 2009; Page et al., 2006) was due to the fact that the earlier studies employed a non-standard Hebb paradigm in which either very long lists were used or/and there was a failure to isolate the possible effects of suppression and phonological similarity specifically on the learning of item-order as opposed to item-identity. It has been argued, therefore, that incidental long-term verbal sequence learning reflects in large part the legacy of the articulatory planning engaged to meet the demands of the short-term recall task, not the action of a phonological store coupled to a separate abstract ordering signal (Sjöblom, 2019; Sjöblom & Hughes, 2020).

The reattribution of key verbal serial short-term recall and learning phenomena ascribed classically to the action of a dedicated phonological store to general-purpose motor and perceptual processes motivates the central goal of the present thesis, namely, to examine the cognitive neuroscience of verbal short-term memory from a perceptual-motor perspective. To date, cognitive neuroscientific research into verbal serial short-term memory has tended to be dominated by the phonological loop model. Indeed, the goal of such research has more often than not been to identify the brain systems that correspond to the individual components of the phonological loop model. The aim of the present research was to examine whether the perceptual-motor account may provide an alternative, more parsimonious, framework within which to understand the cognitive neuroscience of verbal serial short-term memory as well as long-term verbal sequence learning.

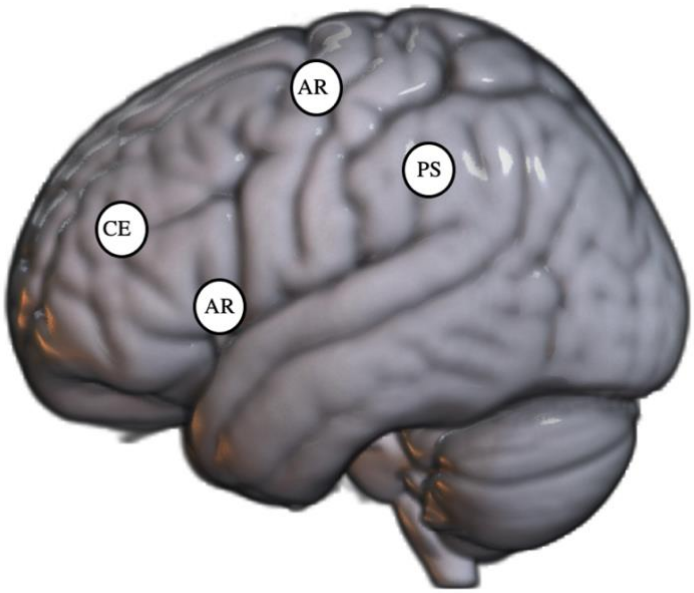
1.5 The Cognitive Neuroscience of Verbal Short-Term Memory Re-Examined

In light of the cognitive-psychological challenges to the phonological loop model discussed in the previous section, the goal of the current section is to begin reassessing the cognitive neuroscience of verbal short-term memory and verbal sequence learning. This seems particularly warranted on the grounds that this area of research, more so, arguably, than even the cognitive psychology of verbal short-term memory and learning, has been dominated by the phonological loop model as its guiding theoretical framework. The phonological store remains perhaps, the most investigated "box" in cognitive neuroscience (Buchsbaum & D'Esposito, 2008). In light of neuroimaging studies claiming to localise the phonological store (to be reviewed below), Baddeley (2003) presented a brain-based diagram of the phonological loop model. Although he may not have intended to remodel phonological short-term memory from 2D boxology to a detailed neurocircuitry, his diagram (see Figure 2) clearly suggests that components of the cognitive model are to be understood as corresponding specifically to particular brain areas: The central executive was proposed to be located in prefrontal cortex, articulatory rehearsal in the motor cortex and inferior frontal gyrus, and the phonological store in the inferior parietal cortex. The store therefore presents itself as an opportunity to use neuroimaging not only for specific localisation but also to challenge a model formulated originally in purely cognitive psychological terms (Buchsbaum & D'Esposito, 2008). When reviewing the early neuroimaging studies throughout this section, it became obvious that the realisation of the phonological loop in the brain was at odds with biological mechanisms as there was little, if any, discussion of anatomical connections from—nor functional properties of—the inferior parietal lobe and its role in learning. This in turn prompted questions about the validity

of the phonological loop concept itself. Instead, the function of the inferior parietal cortex in many cognitive neuroimaging investigations was inferred based on the phonological store construct. Functional neuroimaging should therefore provide the opportunity to examine verbal short-term memory in a new light, where interpretation is guided and constrained strongly by neurobiological considerations as well as cognitive-theoretical accounts.

Figure 2

3D rendering of MNI152 template, adapted from Figure 6 of Baddeley (2003).



Note. The schematic shows the proposed location of the central executive (CE) in the prefrontal cortex, articulatory rehearsal (AR) in the motor cortex and inferior frontal gyrus, and the phonological store (PS) in the inferior parietal lobe.

Prior to the advent and development of functional magnetic resonance imaging (fMRI), neuroimaging investigations throughout the 1990s attempted to localise the phonological store using positron emission topography (PET). Given the progression in experimental techniques and neuroimaging technology, such imaging methods are now deemed inadequate for cognitive neuroscience in any case. Whilst the early PET

studies did initiate the investigations into mapping the phonological loop to anatomical regions in the brain, PET suffers from low spatial resolution which often results in relatively poor anatomical accuracy. Moreover, the early investigations were typically based on a small number of observations (6-10 subjects). Instead, fMRI provides a much more reliable way to investigate haemodynamic activity associated with cognitive functions due to the greater spatial resolution and its ability to assess changes in haemodynamic activity at a relatively high temporal resolution. Being able to assess changes in the BOLD signal in both a spatially and temporally accurate way is of extreme importance because temporal co-activation of regions could provide insight into how specific regions contribute to information processing in the brain.

A first step, then, in evaluating the notion that a particular brain region or network could be identified as the 'substrate' of the phonological store is to identify what functional properties that region/network must exhibit to be compatible with, and to uniquely support, the cognitive-psychological construct of a phonological store (e.g., Buchsbaum & D'Esposito, 2008; Chein & Fiez, 2001, Morey et al., 2019). First, the region/network must be active during passive listening to speech given that auditory-verbal material is proposed to have obligatory access to the phonological store. Second, the region/network must be found outside of any language and speech processing areas as the store is proposed to be functionally and anatomically separate from such regions. The candidate region should also be active during the encoding of to-be-remembered items (as items should be entering the store), during rehearsal (as rehearsal recycles representations through the store to prevent their decay), and finally during recall (when items are being retrieved from the store).

The notion that the phonological store resides in a relatively circumscribed brain region arose from interpretations of neuropsychological data. For

neuropsychological data to constitute clear evidence of a discrete phonological store, however, deficits in storage should be observed in the absence of deficits to language and speech processes, as the store is proposed to be independent from regions involved in such functions. As the majority of patients with substantial temporoparietal lesions have concurrent language and short-term memory deficits (Shallice & Papagno, 2019), those patients should be taken as primary evidence over those suffering from lesions that do not interrupt such functions but are considered to show specific short-term memory deficits (Baddeley & D'Esposito, 2008). These rare cases have shown that they suffer from lesions that in the majority of cases would result in conduction aphasia but by means of pre-existing language abilities and re-organisation potential some are able to recover normal speech production functions. Patient J.B. had a large temporoparietal lesion that included the Sylvian-parietal-temporal (Spt) region, which was very similar to the distribution of lesions associated with conduction aphasia (Buchsbaum et al., 2011; Paulesu et al., 2017). A large body of research indicates that conduction aphasia is caused by damage to the left superior temporal gyrus (STG) and/or the left supramarginal gyrus, regions therefore centred around the temporoparietal area at the posterior portion of the Sylvian fissure (Baldo et al., 2008; Damasio & Damasio, 1980; Green & Howes, 1977). The syndrome results in phonemic paraphasias and impaired verbatim repetition where patients often attempt to self-correct. Whilst deficits in speech production are consistently observed, patients have spared auditory comprehension (Baldo et al., 2008; Benson et al., 1973; Damasio & Damasio, 1980; Goodglass, 1992). Recent work with over 200 stroke patients has further demonstrated that lesions to the left posterior STG cause not only verbal short-term memory impairments but also deficits in speech comprehension, thus supporting

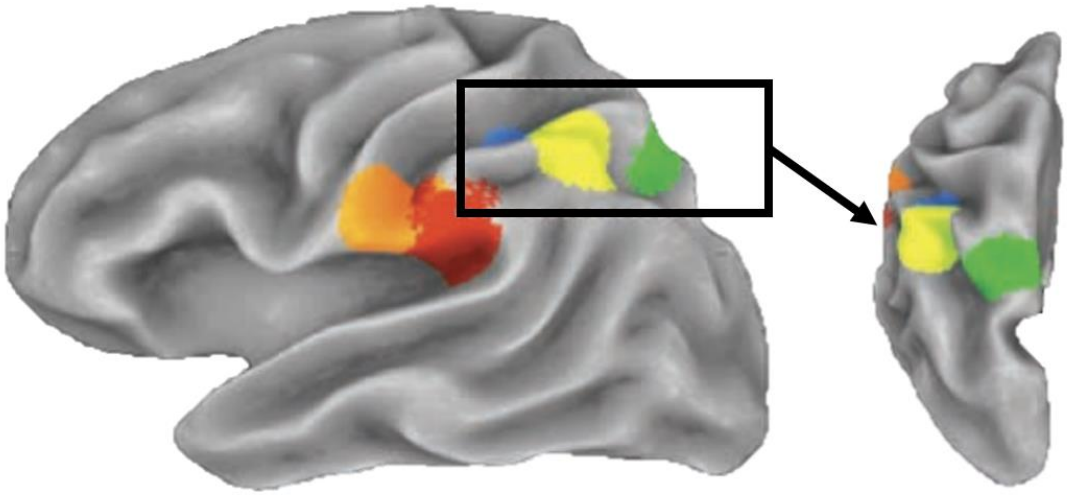
the notion that performance in verbal short-term memory tasks and speech perception/comprehension share the same neural underpinnings (Leff et al., 2009).

It was reported, however, that J.B. maintained her language faculties indicating that she did not suffer from aphasia whilst her verbal short-term memory capabilities were disrupted (Shallice & Butterworth, 1977). Memory deficits were specific to auditory-verbal input and were interpreted as disruption to an "auditory-verbal short-term store" (Warrington & Shallice, 1969). Although such profiles constituted the initial evidence that led to a separation between the articulatory rehearsal process and the passive store (Baddeley & Hitch, 1974) it appears that deficits to J.B's functioning was specific to auditory-verbal input and could not confirm whether deficits also occurred with visual-verbal input. Critically then, deficits should not be taken as evidence for a *phonological* store as this would imply deficits extend to visual-verbal short-term memory tasks also.

Perhaps due in part to the limitations of PET, early neuroimaging studies of verbal short-term memory provided an unclear picture of the location of the phonological store. While there was some consensus that it was located in the parietal lobe, different studies located it in markedly different regions within that lobe (see Figure 3). Depending on the study, activation was revealed in the inferior parietal lobe in the ventral supramarginal gyrus (Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996), more superiorly and posteriorly in the mid supramarginal gyrus (Smith et al., 1996) as well as the posterior parietal cortex (Awh et al., 1996).

Figure 3

Image taken from Buchsbaum and D'Esposito (2008) showing the supposed location of the phonological store (in the left hemisphere) in five PET studies investigating verbal short-term memory.



Note. Co-ordinates were mapped to a cortical surface image in Talairach space. Different colours correspond to different studies: the areas shown in red (Paulesu et al. 1993) and orange (Salmon et al. 1996): ventral supramarginal gyrus; the area shown in blue (Smith et al. 1996): the mid supramarginal gyrus; the areas shown in yellow (Smith et al. 1995) and green (Awh et al. 1996): the posterior parietal cortex.

The posterior parietal lobe, however, can readily be rejected as a plausible location for the store in light of the fact that it is not within the auditory-verbal processing areas. As noted, given that speech is assumed to enjoy direct access to the store, any region/network to be associated with the store should be active even during passive exposure to speech (Becker et al., 1999). Indeed, on the phonological loop model, the disruptive effect of task-irrelevant speech on verbal serial recall is said to occur because the irrelevant speech gains obligatory access to the phonological store (where the to-be-remembered items are also being stored according to this model; Salamé & Baddeley, 1982; but see, e.g., Jones et al., 2004). So, even when speech is not even being listened to (at least not ostensibly), it is assumed that it gains obligatory

access to the store (Baddeley et al., 1984; Baddeley, 2003). But activation is not typically observed in those parietal regions under such conditions (Fiez et al., 1996). Such passive speech processing instead usually activates the primary auditory cortex located in the superior temporal cortex and not areas across the parietal cortex (Binder et al., 2000).

It is worth noting too that the early neuroimaging studies of verbal short-term memory (Awh et al., 1996; Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996) sought to localise the phonological store using the Talairach and Tournoux (1988) stereotaxic co-ordinates which are based only on one elderly post-mortem subject and therefore non-representative of a sample population. The localisation of area 40 to the posterior parietal cortex is considered *too* posterior and is inconsistent with more modern neuroanatomical localisations (e.g., in MNI space). Additionally, non-specific reporting of "areas 40/7" is insufficient. Such reporting suggests that authors were unable to clearly localise clusters of activation to one specific area. If activation clusters bordered the two areas, then authors should have highlighted this in their discussion of results by referring to sulcal and gyral anatomy- this however was rarely the case. The term "areas 40/7" is nonetheless used in Table 1 to highlight inconsistencies across investigations. Gross anatomy and the areas stated in the forthcoming review are reported as they were in the original studies. Table 1 seeks to highlight the differences in parietal lobe activation across verbal short-term memory tasks.

A further difficulty in accepting the localisations of a phonological store to the inferior parietal lobe (area 40) in the early neuroimaging studies is that those localisations are more superior than areas typically associated with short-term memory deficits in patients with focal brain lesions (Shallice & Warrington, 1980; Vallar &

Papagno, 1995; Shallice & Vallar, 1990). Some studies, however, did not observe activation in the parietal lobe at all (Fiez et al., 1996; Grasby et al., 1993). And in others, the inferior parietal lobe was active but only during certain phases of a verbal serial recall task (Chein & Fiez, 2001). This is highly problematic for the phonological store theory because, as noted, activation of the region(s) to be identified with the phonological store should be observed across all task phases (Chein & Fiez, 2001; Buchsbaum & D'Esposito 2008; Morey et al., 2019).

Table 1

Brain regions found to be active during tests of verbal short-term memory. Brodmann areas indicated by an asterisk (*). All studies were localised using Talaraich and Tournoux (1988) co-ordinates other than those localised in MNI space indicated by (^). Cerebellar lobules as reported by Desmond et al. (1997) are indicated by (°) as many studies did not localise to specific cerebellar lobules. If reported in "lobule unspecified" column the original paper nor Desmond et al., (1997) reported the specific lobule. 'L' refers to left hemisphere, 'R' to right hemisphere, 'M' to midline and 'B' bilaterally in both hemispheres.

	Brain area or lobule	46	9	10	6/9	6	44	45	22	42	7/40	7	40	Cerebellum (lobule unspecified)	IV/V	VI	VIIA	VIIB	VIII
Task																			
Awh et al. (1996)																			
Verbal 2-back						L	L					L	L	R					
2-back – search control						B	L					B	L	B					
2-back – rehearsal control						B						B	L			M°		R°	
Chein & Fiez (2001)																			
Presentation						L*	L*	L*			L*								

	Brain area or lobule	46	9	10	6/9	6	44	45	22	42	7/40	7	40	Cerebellum (lobule unspecified)	IV/V	VI	VIIA	VIIB	VIII
Chein & Fiez (2001)																			
Maintenance						L*	L*	L*											
Recall						L*		L*	B*	B*									
Chen & Desmond (2005a)																			
Memory task		R	R	R		L	L	R					B			B	B Crus I	R	
Rehearsal control						L	L									B	B Crus I		
Chen & Desmond (2005b)																			
Encoding						B	L		R						R	B	B Crus I	R	R
Maintenance							L	B					L					R	R
Retrieval						L	L	R								L			

	Brain area or lobule	46	9	10	6/9	6	44	45	22	42	7/40	7	40	Cerebellum (lobule unspecified)	IV/V	VI	VIIA	VIIB	VIII
Desmond et al. (1997)																			
Memory task																R	L	R	
Rehearsal control																R	L		
Henson et al. (2000)																			
Letter match – Symbol match			L*			L*	L*					B*	B*						
Letter probe – Letter match		R*		B*									B*						
Sequence probe – Letter probe			L*																
Sequence Probe – Grouped probe						L*													
Paulesu et al. (1993)																			
Item recognition (phonological memory)						B*	B*						B*	B					

	Brain area or lobule	46	9	10	6/9	6	44	45	22	42	7/40	7	40	Cerebellum (lobule unspecified)	IV/V	VI	VIIA	VIIB	VIII
Salmon et al. (1996)																			
Item recognition (phonological memory)					L*	L*						L*							
Sakai et al. (2002a)																			
Presentation																			
Maintenance		B*^				B*^	L*^					B*^				B*^			
Recall		B*^				R*^	L*^												
Smith et al (1996)																			
Verbal match judgement (3-back)	B	B	B			L	L			L	R								
Spatial match judgement (3-back)	B	R	L		B					L	B	B							
Verbal 2-back					B	L					B	L	B						

The majority of the early PET studies used the Sternberg (1966) task to assess verbal short-term memory. This task comprises three phases: an 'encoding' phase, where letters are visually presented to the subject, a 'maintenance' phase, where subjects are instructed to retain the stimuli over a brief delay, and a 'retrieval' phase, where subjects are presented with a probe letter and have to respond as to whether or not the letter was present in the just-presented array. Whilst it appears that such a task would be suitable for testing verbal short-term memory, the majority of early PET studies did not specify which phases of the memory task were contrasted with control tasks, or whether a comparison of different phases was considered at all. Although authors assumed that the phonological store should be active regardless of task phase, it is likely that activation interpreted as the location of the store was only observed in one or two of the task phases but appeared in results of whole trial comparisons. Moreover, there is evidence that this type of task is more likely to depend on recognition, based predominantly on a retrospective strategy (e.g., familiarity; Lovett et al., 1999) whereas serial recall tasks are typically thought to require a more proactive rehearsal strategy (Durisko & Fiez, 2010).

The PET studies of Paulesu et al. (1993) and Salmon et al. (1996) both used the Sternberg (1966) task. Alongside this 'phonological memory' task, they also used a visual short-term memory control task that was identical except that the letters were Korean, which the authors assumed could not be transcoded into a phonological form. When neural activations associated with the phonological memory task were compared with those from the Korean control task, activation was found bilaterally in the supplementary motor area (SMA; BA 6), Broca's area (BA 44), superior temporal gyrus (BA 22/42), supramarginal gyrus (BA 40) and the cerebellum. The supramarginal gyrus (SMG) was identified as the primary location of the phonological

store while the activations in BA 44, BA 22 and BA 42 were taken to reflect the articulatory rehearsal process. Although the authors address the involvement of the SMA (BA 6) and cerebellum in motor aspects of speech planning (despite the absence of overt speech), somewhat oddly they did not implicate these regions in articulatory rehearsal. In a further attempt to localise other specific components of the phonological loop model, Paulesu et al. (1993) aimed to separate the functional anatomy of subvocal articulatory rehearsal from the phonological store by comparing a rhyming judgement task with the item recognition (phonological memory) task. They assumed that the phonological memory task would engage both articulatory rehearsal and the phonological store while the rhyming judgement task would only activate regions involved in articulatory rehearsal. Subjects were shown a sequence of letters and had to state whether or not each letter rhymed with the letter 'B' which was always present on the screen. Activation in the left supramarginal gyrus (BA 40) was found in the phonological memory (item-recognition) task but not the rhyming judgement task. The authors inferred that the absence of supramarginal activation in the rhyming judgement task and proximity of their activation to lesion sites in patients who show verbal short-term memory deficits (McCarthy & Warrington, 1990; Shallice & Vallar, 1990) meant they had localised the phonological store.

The conclusions of Paulesu et al. (1993) are problematic for multiple reasons, however, from the standpoint of the phonological loop model. First, rhyming decisions were said to engage subvocal rehearsal but not the phonological store (Burani et al., 1991; Vallar & Baddeley, 1984). The assumption that rhyme judgements do not engage the phonological store rests on patient J.B. having a defective phonological store but preserved ability to perform rhyme judgements (Vallar & Baddeley, 1984). Therefore, using the rhyme judgement task to probe and locate the phonological store

in an imaging study is contingent on having correctly linked the phonological store (or absence of it) to a particular lesion in a single patient (Buchsbaum & D'Esposito, 2008). This is particularly problematic as a review of neuropsychological data has shown that previous interpretations of a deficit to phonological storage was in fact specific to auditory-verbal material and cannot be generalised to visual-verbal input. Moreover, basing interpretations on the symptom profile in a single patient is non-representative of population. Second, according to the phonological loop model, the act of articulatory rehearsal grants graphemic input access to the phonological store and also refreshes decay-prone representations already in the store. Thus, if articulatory rehearsal is required for rhyme judgments on visually presented items as Paulesu et al. (1993) assumed—and for which there is indeed independent evidence (e.g., Besner, 1987; Besner et al., 1981; Tree et al., 2011)—then on the phonological loop model the phonological store must also be implicated in the task because articulatory rehearsal inevitably activates the phonological store. Their logic, therefore, appears to be at odds with the cognitive model on which their interpretation of the neural evidence is based.

Other neuroimaging evidence also suggests that the interpretation of a discrete phonological store in the supramarginal gyrus is particularly problematic. The SMG maintains reciprocal connections to the ventral premotor cortex and IFG (pars opercularis) regions typically associated with articulatory motor planning (Catani et al., 2005; Petrides & Pandya, 2009; Rushworth et al., 2006). This indicates that the SMG is not an anatomically discrete region separate from motor or perceptual regions involved in speech and language. Additionally, some have suggested that activation of the SMG is related to stimulus encoding rather than mechanisms directly related to retention or recall (Fiez et al., 1996b). In line with this view, results of a meta-analysis

of 35 neuroimaging studies of reading indicated that the supramarginal gyrus is implicated in a network of regions for reading involving left lateralised superior temporal areas, and the IFG (pars opercularis) (Jobard et al., 2003). Whilst the SMG is implicated in "phonologically" demanding tasks, activation of the gyrus during rhyme (Petersen et al., 1988), syllable (Price et al., 1997; Devlin et al., 2003) and phoneme judgments (Zevin & McCandliss, 2005; Raizada & Poldrack, 2007) has been argued to be due more to the articulatory requirement of those tasks rather than the involvement of the region in storing abstract verbal representations (Pattamadilok et al. 2010). Lastly, the SMG has been shown to contribute to visually guided hand actions (Binkofski et al., 2004; Rushworth et al., 2001; Price, 2010). Overall, these implications demonstrate that the SMG is not specific to the retention of abstract verbal representations as suggested by the phonological loop model.

Another attempt to localise the phonological store was predicated on the assumption that separate cognitive sub-systems exist for verbal and spatial short-term memory as described in the Working Memory Model (Baddeley & Hitch, 1974). Smith et al. (1996) used verbal and spatial 3-back tasks in an attempt to parse the differences between neural activity during the two types of short-term memory task. Subjects were required to indicate whether the item identity or location of the item on the screen was the same as the item presented three items earlier (hence '3-back'). Results showed that the same region—one bordering area 7 and 40 in the left mid supramarginal gyrus—was identified in both tasks suggesting that the region is not functionally specialised for verbal material. Moreover, activation of left area 40 in a verbal 2-back task (with no spatial comparison) was in very close proximity to an activation cluster identified in left area 40 in the spatial 3-back task. It seems implausible to suppose that two regions separated only by a matter of millimetres would constitute the neural

substrates of what are meant to be functionally, highly distinct modules. Additionally, the activation identified in the mid supramarginal gyrus was more superior and posterior to the localisations of Paulesu et al. (1993) and Salmon et al. (1996); thus, no common region has been identified across a number of studies all of which were designed to localise the phonological store (see Figure 3).

It appears that phonological store-based narrative continued to greatly influence interpretations in other neuroimaging investigations. Henson et al. (2000) proposed that the activation of BA 40 in the left SMG was activated by demands related to phonological storage when subtracting activation in a letter match task (judging whether upper- and lower-case letters presented together were the same) from a letter probe task where subjects were presented with a single letter and required to indicate whether it had appeared in the just-presented six-letter sequence. However, activation of BA 40 was also observed in the results of a letter match—symbol match subtraction which was meant to isolate regions involved in the recoding of visual-verbal stimuli. On the phonological store theory, the correlate of the phonological store should also be observed during such visual-verbal recoding (because representations would be entering the store). Yet the activation of BA 40 (reported as the inferior parietal gyrus) in the results of the letter match—symbol match contrast was not discussed in terms of the possible contribution of the area to phonological storage. The fact that activations of BA 40 along the SMG and inferior parietal gyrus were not similarly discussed with reference to phonological storage calls into question the attribution of phonological storage to BA 40 at all. Results appear to have been overly influenced by others claiming to have already localised the store to a similar location[1]

[1] The authors stated that their localisation was in line with Awh et al. (1996) and Paulesu et al. (1993). While it was in line with Paulesu et a. (1993), the left supramarginal gyrus was not in fact identified as the locus of the phonological store by Awh et al. (1996; see Figure 3).

and further highlight those interpretations of results from neuroimaging investigations into verbal short-term memory often are at odds with phonological store theory.

It also appears that the localisation of the phonological store to the inferior parietal gyrus was misreported as gyri in the inferior parietal *lobe* are usually referred to as supramarginal or angular. If the authors were instead referring to the intraparietal sulcus (as activation associated with the letter match task was more superior to that associated with the letter probe task) then the activation cluster was in close proximity to regions along the intraparietal sulcus often implicated in visual processing, again undermining the attempt to associate this region exclusively with phonological storage.

Aside from attempting to localise the phonological store, Henson et al. (2000) also sought to test the notion discussed earlier in Section 1.3 that articulatory rehearsal—assumed on the phonological loop to maintain phonological item-traces—can be dissociated from a timing signal that represents item-order (cf. Burgess & Hitch, 2005). A sequence probe task used individual presentation of memory items, but the probe consisted of all six letters presented simultaneously. Subjects had to judge whether the entire sequence was for a match of the serial order. A grouped probe task was identical to the sequence probe task except that a short pause was inserted after every third letter, temporally grouping the items into two groups. Results did not demonstrate the predicted dissociation in brain activations however: the premotor cortex was found to be active both in a subtraction designed to reveal the location of the timing signal (grouped probe—sequence probe) and in a subtraction designed to reveal the location of the articulatory rehearsal process (sequence probe—letter probe). As activation of the pre-motor cortex was observed in the *sequence probe—letter probe* contrast and the *grouped probe—sequence probe* contrast, the authors appeared

to suggest that the articulatory rehearsal process and the timing signal may not in fact be separate components.

Not only are the neuroimaging data reviewed thus far incompatible with the cognitive model of phonological storage, but a good deal of heterogeneity in terms of attempts to localise the store is evident. In contrast, all studies have associated activity (in varying combinations) of areas 6, 44 and the cerebellum with an articulatory rehearsal process (see Table 1). Some interpretations of verbal short-term memory, however, *do not* claim to have localised a phonological store to a region in the parietal cortex but do report activation of the same regions associated with articulatory rehearsal. This places a greater emphasis on the operation of an articulatory rehearsal process and further informs the predictions of the perceptual-motor approach. A key contention of the present thesis then, is observing whether and how patterns of activation differ as a function of task phase is critical for evaluating whether there is neuroscientific evidence for a phonological store, and for adjudicating between the phonological store theory and the perceptual-motor account.

In an attempt to isolate brain regions involved in encoding (of visual-verbal to-be-remembered items), rehearsal, and recall, Chein and Fiez's (2001) fMRI study used a delayed serial recall task using sets of verbal items in which phonological similarity, syllable length and lexical status were manipulated. Following visual presentation of the stimuli and a retention interval, subjects were required to recall the word lists out loud. The central tenet of Chein and Fiez's (2001) argument was that if a store were to reside somewhere in the cortex, the associated region should remain active throughout all task phases. Critically, however, the left inferior parietal lobe (BA 40/7)—which had previously been associated with the phonological store (Awh et al., 1996; Smith et al., 1996)—was only shown to be active during presentation and

was not observed during the retention interval (regardless of stimulus manipulation) whilst the premotor cortex (BA 6) and IFG (BA 45) were found to be active throughout all task phases. The opercular part of IFG (BA 44) was active during both presentation and retention interval whilst the prefrontal cortex (BA 46) was only observed during the retention interval (see also Chein & Fiez, 2010; Fiez et al., 1996). In contrast, other attempts at isolating activation specific to task phases in a verbal short-term memory task only observed activation of the intraparietal sulcus across the delay period and not during presentation whereas pre-frontal cortex (BA 46), pre-motor and SMA (BA 6) and inferior frontal gyrus (BA 44) were active across the delay and the response stage when presented with a probe (Sakai et al., 2002a).

Studies that are likely to promote the covert articulatory rehearsal of visually presented items or that require a match-judgement response have consistently shown activation in the left pre-motor cortex (BA 6), posterior inferior frontal gyrus (BA 44) and areas in the cerebellum (Awh et al., 1996; Chein & Fiez, 2001; Henson et al., 2000; Jonides et al., 1998, Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996). Activation of such regions is also observed in tasks comparing verbal maintenance to silent counting (Fiez et al., 1996), requiring serial repetition (Becker et al., 1994) and free recall (Becker et al., 1994; Grasby et al., 1993), thus suggesting a fundamental role for these areas in verbal serial short-term memory performance.

Indeed, an increasing body of evidence highlights the similarities between activations associated with verbal short-term memory phenomena and perceptual-motor processing in regions involved in speech and language abilities (Buchsbaum et al., 2011; Jacquemot & Scott, 2006; Buchsbaum & D'Esposito, 2008; Fegen et al., 2015; Koenigs et al., 2011). Following on from the foregoing consideration of the neuroimaging literature, the next section seeks to evaluate the extent to which the

cognitive neuroscience of verbal short-term memory can be reconceptualised in vocal-articulatory and perceptual processing terms. In particular, the lack of consistency in attempts to localise a phonological store in a discrete brain region separate from motor and perceptual processing regions motivates the re-examination of the cognitive neuroscience of verbal serial short-term memory from a perceptual-motor perspective.

1.6 Can the Neural Basis of Verbal Short-Term Memory be re-explained in perceptual—motor terms?

In light of the recent behavioural evidence suggesting that verbal serial short-term memory performance may be explained parsimoniously in terms of articulatory and perceptual processes (Hughes & Marsh, 2017; Hughes et al., 2009, 2016; Jones et al., 2006, 2004; Macken et al., 2016; Sjöblom & Hughes, 2020; see Section 1.4), the current section seeks to outline how such performance may, in neuroscientific terms, be readily explicable in terms of the operation of brain regions involved in a variety of functions rather than a singular region dedicated solely to the temporary storage of modality-independent (phonological representations)

The hierarchical organisation of the frontal lobe suggests a way in which to begin mapping a perceptual-motor approach to verbal short-term memory onto a neural level of analysis. A general view is that the hierarchy of the frontal cortex constitutes a production-system architecture, where the role of neurons in the prefrontal cortex is to represent context information related to integrated representations of task rules, instructions and temporal contingencies necessary to support an appropriate behavioural response. This context information is essential for mediating actions separated in time but contingent on one another such that they comprise a sequence of fluid actions (Fuster, 1990; Miller & Cohen, 2001). One view

is that human language and speech motor control evolved from manual gestures (cf. Hewes, 1973) and is made possible through the diverse and flexible way humans can represent information and communicate that information using their bodies. It therefore seems reasonable to assume that the planning and reproduction of a verbal utterance occurs in the same way as other motor behaviours (Jones & Macken, 2018).

Action-related goals maintained in the prefrontal cortex (PFC) are encoded in abstract terms (Cohen & Servan-Schreiber, 1992; Passingham, 1996). This means that while a goal (e.g., to reproduce a sequence) does not specify the order of actions required to achieve that goal, or even the effector required to execute the action, the goal itself can be broken down into particular sub-goals until they become specific motor actions (such as the articulation of individual words or other verbal items). Verbal serial short-term memory performance could, therefore, be achieved via the maintenance of abstract representations regarding perceptual input and forthcoming operation as the PFC is proposed to facilitate the activation of networks involved in the reception of sensory signals and in the execution of motor actions that are part of a sequence (Fuster, 2000; Miller & Cohen, 2001). The mediation of cross-temporal contingencies (between perception and action) would appear to map well onto the notion of temporal binding assumed in the perceptual-motor account where vocal-motor planning in verbal short-term memory tasks is designed to sequentially bind items that are otherwise unrelated (e.g., Hughes et al., 2009). Although neuroimaging studies of verbal short-term memory have consistently localised articulatory rehearsal to the premotor cortex and/or the SMA (see Table 1), such mapping has only ever been considered in terms of the role of articulatory processes in supporting a distinct phonological store (Awh et al., 1996; Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996). On the perceptual-motor account, in contrast,

articulatory rehearsal has been reappraised as the very method by which a verbal sequence is retained and reproduced. The lateral PFC (area 46) is the apex of the production-system hierarchy, whilst the premotor and primary motor areas constitute the lower levels (Fuster, 2004). As the premotor cortex (BA 6) receives significant projections from area 46[2] (Orioli & Strick, 1989; Lu et al., 1994), it is widely accepted that the PFC supports action-planning via the selection and maintenance of goals at higher levels (e.g., 'reproduce a sequence') right through to the motor responses involved in the planning or execution of the action (Badre & D'Esposito, 2009; Lashley, 1951; Miller et al., 1960).

From the perspective of the perceptual-motor account of verbal serial short-term memory, it could be suggested that abstract goal and rule related information in the PFC is translated into motor commands at a lower level of abstraction in the premotor cortex (see rostral-caudal gradient; Badre & D'Esposito, 2009). As the to-be-remembered items are presented in an unfamiliar order, some form of short-term processing is necessitated "in the moment." The sequential and temporally structured nature of articulatory planning is therefore co-opted to turn the succession of discrete, sequentially unrelated, items into a singular "motor object" (Sternberg et al., 1980). An example of the way in which relatively discrete items as-presented are temporally bound into a more singular motor-object is lenition Typically, entire vowel sounds are dropped (Bauer, 2008) thus reducing the amount of information that needs to be bound

[2] It should be noted that there are differences in the localisation of area 46, BA 9, and area 9/46. The central part of the middle frontal gyrus has been described as area 46 (Rajkowska & Goldman-Rakic, 1995), whereas Brodmann (1909) referred to this location as area 9 (BA 9). Petrides and Pandya (1999, 2002) reported that BA 9 has cytoarchitectonic properties similar to those of classical area 46, and so referred to it as area 9/46. Each are proposed to encode abstract information such as rules (Freedman et al., 2001; Wallis et al. 2001).

into a singular verbal object during normal speech. Examples include the pronunciation of ‘every’ and ‘evry’ and ‘memory’ as ‘memry.’

Thus, whilst representations of sensory features and prospective action may be represented in abstract form, it is suggested, on the basis of the perceptual-motor account, that it is the motoric-level sequencing of those representations in the premotor cortex that maintains and indeed *generates* their order. It is widely accepted that the premotor cortex is involved in the planning and generation of action commands (Weinrich & Wise, 1982; Wise, 1985). The region is a direct recipient of commands from the PFC (Orioli & Strick, 1989; Lu et al., 1994), is active prior to action execution and subsequently projects directly to the primary motor cortex (Lu et al., 1994) and spinal cord (Dum & Strick, 1991). Indeed, the abstract representations of goals in the PFC may be epiphenomenal to meeting the demands of the memory task. That is, in this view, the motor sequencing is not in the service of keeping abstract (e.g., phonological) representations active in a separate store dedicated to holding such representations but to generate a new sequential (motor) object that will allow for the reproduction of the order of the presented items. In particular, regions of the ventral premotor cortex and IFG (pars opercularis) are proposed to generate articulatory codes that are then used by the premotor and primary motor cortex for speech planning and production. Articulatory rehearsal has consistently been associated with the left premotor cortex and IFG in verbal short-term memory tasks (Awh et al., 1996; Chein & Fiez, 2001; Fiez et al., 1996; Paulesu et al., 1993; Salmon et al., 1996). Tellingly, these regions have also been found to be active in other tasks requiring speech planning or production (Callan et al., 2010; Hickok, 2009; Hickok et al., 2011; Meister et al., 2009; Papoutsi et al., 2009; Rogalsky et al., 2008). This activation is proposed to reflect the mapping of intended speech to orofacial movements (Price, 2012).

Whilst activation in the left premotor cortex and IFG during verbal serial short-term memory tasks was consistently attributed to an articulatory rehearsal process (Awh et al., 1996; Fiez et al., 1996; Paulesu et al., 1993; Salmon et al., 1996), interpretations often failed to consider how perceptual input would be mapped onto the motor plan and what regions of the brain this may involve. The PFC is proposed to facilitate the activation of networks involved in receiving perceptual signals and networks necessary for the planning and execution of motor actions (Fuster, 1997). It would appear then that abstract representations of rules and goals in the PFC could mediate the perceptual-motor mapping process described in the perceptual-motor account (e.g., Hughes et al., 2009; Macken et al., 2016). From a phonological-store based view, auditory-verbal information is proposed to have direct and obligatory access to the phonological store (because it is deemed to already be in a suitable code) while visual-verbal information in contrast needs to be converted into the store's code indirectly via a grapheme-phoneme conversion process carried out by the articulatory control process (Baddeley, 2003, see Section 1.5, Figure 2). From a phonological store perspective, then, auditory-verbal input would not necessarily need to be mapped onto a motor plan. But theories of speech perception and production suggest otherwise.

In the auditory domain, perceptual-motor mapping or sensorimotor integration is proposed to be achieved via feed-forward and feed-back projections supported by the arcuate, middle and superior longitudinal fasciculi (Saura et al., 2008; Ueno et al., 2011). These enable, respectively, predictive motor signals to modify activity in the superior temporal auditory cortex for perceptual tuning and predictive sensory signals to modify activity in premotor and motor cortices tuning production (Hickok & Poeppel, 2004, 2007; Hickok et al., 2011 Hickok, 2012; Rauschecker and Scott, 2009; Rauschecker, 2011). Specifically, once a speech signal is decoded in the superior

temporal cortex, representations are suggested to be transformed into to motor-articulatory representations in the IFG (pars opercularis; BA 44) and ventral premotor cortex (BA 6) (Rauschecker & Scott, 2009).

A critical region within the auditory-motor network is the Sylvian-parieto-temporal (Spt) region. This region has been suggested as a candidate region for the location of the phonological store based on neuropsychological data (Shallice & Butterworth, 1977; Warrington and Shallice, 1969). This idea was discounted, however, due to its location *within* the auditory-verbal zone and more recent evidence implicating the Spt in a variety of perception and production processes suggests that it is not exclusive to memory and thus could not be the location of the store. Moreover, activation of the Spt is correlated with activity in frontal areas related to speech production such as pars opercularis (BA 44) suggesting that the two are functionally connected (Buchsbaum et al., 2001; Buchsbaum et al., 2005b). Using time-frequency analysis, Herman et al. (2013) showed the rise and fall of oscillatory power that alternates between the IFG and Spt, consistent with a coordinated exchange of information. This indicates that the Spt is motor-effector selective, showing greater activation when the output task involves the vocal tract compared with the manual effectors (Pa & Hickok, 2008). Damage to a sensory-motor integration circuit, therefore, can lead to an impairment in the capacity for auditory representations of speech to guide and constrain corresponding articulatory representations thought to be generated in the inferior frontal gyrus and ventral premotor cortex (Hickok & Poeppel, 2000, 2004; Wise et al., 2001). The Spt is therefore argued to act as an auditory-motor interface that is integral to speech perception and production.

The region has shown to be active in verbal short-term memory tasks across extended delay periods during covert maintenance as well as during presentation of

stimuli (Buchsbaum & D'Esposito, 2008; McGettigan et al., 2011). The region is also active during speech production tasks that require little to no memory capacity such as single word reading (Buchsbaum et al., 2005b), pseudoword learning (Graves et al., 2008) and other tasks requiring covert (Buchsbaum et al., 2001; Buchsbaum et al., 2005b; Fegen et al., 2015; Hickok et al., 2003; Okada et al., 2018; Wildgruber et al.,1999) and overt speech (Shuster & Lemieux, 2005). It appears then that rehearsal of verbal information is achieved by feed-forward and feed-back pathways connecting the auditory and motor systems (Buchsbaum & D'Esposito, 2008). The perceptual-motor mapping between auditory information and a motor plan assumed within the presently adopted account of verbal serial short-term memory may therefore emerge from flexible interactions within a fronto-temporal sensorimotor cycle that has evolved to support the perception and production of speech (Buchsbaum & D'Esposito, 2019).

It may be, then, that similar operations occur for the perceptual-motor mapping of visual stimuli via a fronto-occipital sensorimotor cycle. Although the integration of visual-verbal information with motor information has not been researched as widely, an attempt is made here to highlight which regions may be involved in the perceptual-motor mapping of visual-verbal stimuli. Within the ventral occipitotemporal cortex (vOTc), the fusiform gyrus is proposed to hold the visual word form area (VWFA) and has consistently been shown to be involved in the processing of single letters and words (Cohen et al., 2000, 2002; Wandell, 2011). Some suggest that much like auditory processing, orthographic processing is also incorporated into the speech production architecture (Carreiras et al., 2014). Anatomical connections via the superior longitudinal and inferior fronto-occipital fasciculus may enable the transformation of orthographic form to articulatory form (Bernal & Altman, 2010; Wakana et al., 2004; Yeatman et al., 2013). A functional coupling is suggested to exist

between the left occipito-temporal cortex (vOTc) and inferior frontal regions for phonological processing (Bokde et al., 2001). Similar to the oscillatory neural patterns observed with the IFG and Spt, activation in the IFG (pars opercularis) has been shown to peak at the same time as that in the left vOTc during visual word processing (Cornelissen et al., 2009). More specifically, activation of the left vOTc and IFG (pars opercularis) has been attributed to the translation of visual forms to articulatory representations (Purcell et al., 2011; Rapp & Dufor, 2011; Rapp & Lipka, 2011) and are consistent with anatomical connections via ventral pathway departing from the occipito-temporal cortex reaching limbic structures in the temporal (Turner et al., 1980) and ventral frontal lobe (Kuypers et al., 1965). Whilst neuroanatomical models of reading have also implicated the supramarginal gyrus (SMG) and regions across superior temporal cortex (STc) (Jobard et al., 2003) the approach taken in this thesis will be to focus solely on the vOTc and motoric regions because the roles that these regions may play in perceptual-motor mapping can be more readily envisioned. Moreover, the activation of the SMG and STc has been shown to vary considerably depending on the familiarity of verbal stimuli, consistency of orthography to phonology and the task requirements.

Based on the current approach, it is hypothesised that the IFG (pars opercularis) performs the same function following auditory or verbal input which is then apprehended by the premotor cortex. Critically, different regions involved in this process are predicted to be implicated depending on presentation modality (Spt and vOTc for auditory and visual stimuli respectively). However, it is not only the operations of regions across the cerebral cortex that are proposed to enable an articulatory planning strategy for verbal serial recall. Neuroimaging investigations into verbal serial-short term memory often overlooked the contributions of different sub-

regions across the cerebellar cortex. Although activation of the cerebellum was reported (Awh et al., 1996; Paulesu et al., 1993, cf. Table 1), little—if any— detailed consideration was given as to how cerebellar cortical circuitry could support the retention and reproduction of a sequence. Neuroanatomical theories of motor skill learning which implicate the cerebellum maintain several parallels with the perceptual-motor approach to verbal serial recall and sequence learning. The cerebellum therefore presents itself as a likely candidate to facilitate long-term verbal sequence learning as a specific instance of motor skill learning where such learning is supported by the repetition and increasing fluency of an articulatory plan.

The cellular organisation of the cerebellum has inspired its inclusion in a number of theories of learning. Marr (1969), the earliest and most influential theorist on this topic, suggested that cerebellar plasticity is achieved through the strengthening of synapses between the dendrites of Purkinje cells (the principal computational unit of the cerebellar cortex) and one of its principal inputs, parallel fibres. It was suggested that this takes place through Hebbian principles, such that the co-activation of these components increases the efficacy of this particular synapse. Crucially, the process depends upon an error signal conveyed by the other principle input to Purkinje cells, climbing fibres[3], in the form of "complex spikes" and is thought to convey discrepancies between actual and intended movements. Decreasing errors would result in the decrease in the frequency of complex spikes, and such support is available from neurophysiological data from behaving animals. Recordings from Purkinje cells from cerebellar lobules III, IV and V of non-human primates have found a high frequency of spike activity at the start of learning, which declined to background levels as

[3] This error signal does not indicate a motor error but instead reflects unexpected sensory events (Kawato & Gomi, 1992; Porrill et al., 2004)

learning progressed (Gilbert & Thach, 1977). Building on this work, Albus (1971) suggested that information storage through these mechanisms was more likely to be achieved through a trial-by-trial progressive *decrease* in the efficacy of parallel fibre - Purkinje cell synapses. Over time, Purkinje cells may learn to respond to signals from parallel fibres under the guidance of a teaching signal from climbing fibres (Ramnani, 2006). Evidence in support of this mechanism includes the finding that the efficacy of Purkinje cell-parallel fibre synapses can be decreased using experience-dependent Hebbian mechanisms *in vitro* (Long-term depression, LTD). Long term potentiation (LTP) has also been observed in cerebellar circuits when parallel fibres are stimulated independently of climbing fibres (Crepel & Jaillard, 1991; Salin et al., 1996). This suggests that parallel fibre-Purkinje cell plasticity is bi-directional and that both mechanisms must occur for the cerebellum to act as a retention and learning device (Ito, 2006). The models (neural representations) acquired through learning can simulate natural processes such as movement (Wolpert & Miall, 1996) where learning and automaticity are reflected in decreases in cerebellar activity (Balsters & Ramnani, 2011; Doyon et al., 2002; Imamizu et al., 2000; Penhune & Doyon, 2005). However other research has shown an increase in excitability in the cerebellum when learning constant timings of finger movement sequences (Jueptner et al., 1997; Ramnani & Passingham, 2001).

Control theoretic models also form a framework to explain how feedback mechanisms are used to optimise the control of movement and decision-making. It has been argued that they might also explain the operations of cerebellar circuitry and its interactions with connected brain areas (Ramnani, 2014). In control theory, there are two types of internal model (Kawato & Wolpert, 1998; Wolpert & Kawato, 1998; Wolpert & Miall, 1996). A “forward” model involves the control of a “plant” and in

the context of the motor apparatus required to generate speech this would include the control of the vocal tract, larynx, tongue and lips. A motor command from the motor cortex, on its way to the muscles, would be copied to a system that uses it to predict the sensory consequences of movements were the command to be executed ("efference copy"). That prediction would be compared with the actual outcomes. The discrepancy between predicted and actual outcomes represents a prediction error that drives the system to learn in ways that minimises this error. It has been suggested that forward models used to generate predictions are instantiated in cerebellar circuitry using mechanisms similar to those described above and that, similarly, error signals that modify these forward models arrive via climbing fibre inputs. This view is supported by extensive evidence that cerebellar circuitry is capable of plasticity (Boyden et al., 2004), that fibres conveying motor commands to muscle groups branch, and that these branches carry copies of motor commands to the cerebellar cortex (Ugolini and Kuypers). Over time, error-signals are reduced as predictions of the forward-models becomes increasingly accurate. Learning, therefore, is a transition from controlled processing where performance is flexible and responsive to ongoing error feedback, to a state where behaviours or movements become efficient, feed-forward and automatic - dependent on prior experience rather than ongoing feedback (Ramnani, 2014).

Which areas might generate information that is copied to the cerebellum? Ventral areas of the precentral cortex are activated by speech (Brown et al., 2008), and these are likely to send outputs to musculature that controls speech. It is also likely that motor commands arrive at cerebellar circuitry via the pontine nuclei. However, as explained above, the motor cortex is part of a hierarchically organised neocortical system in which the abstract goals of action are elaborated into contextual motor

commands. The prefrontal cortex can itself therefore act as a controller, sending commands to the premotor cortex for motor planning or to regions in the temporoparietal cortex for perception (Ito, 2008; Ramnani 2014). Forward models could, then, play roles in the long-term learning for verbal sequences. As discussed previously, area 46 in the prefrontal cortex sits at the top of a hierarchically organised control system for action planning and execution given its direct connectivity with the premotor cortex (Passingham & Wise, 2012). On this logic, the prefrontal and premotor cortices could be involved in verbal short-term memory to the extent that they are responsible for the articulatory planning necessary to reproduce a verbal sequence but also because they receive information from forward-models acquired in the cerebellum necessary for error adjustment and successful serial recall.

Traditional accounts of cerebellar information processing integrate cerebellar anatomy, physiology and theory into explanations of how cerebellar circuits might acquire and "store" motor memory. As the cerebellum has classically been viewed as a motor structure, the early literature focussed on cerebellar involvement in motoric aspects of human behaviour and suggest that the process of motor learning enables motor engrams to be acquired within the plastic circuits of the cerebellar cortex (Ito, 2000, 2002, 2005; Marr 1969). Once the representation is fully acquired, it can be used for the automatic execution of actions without as much involvement from cerebral cortical structures. Thus, the use of internal models (Kawato & Wolpert, 1998; Miall & Wolpert 1996; Wolpert & Kawato, 1998) releases cortical information processing capacity for resolving novel problems. In support of this view, studies have documented dynamic changes in cerebellar activity that accompany motor learning (Jueptner et al., 1997; Passingham, 1996; Ramnani & Passingham, 2001).

In a challenge to the traditional motor view of the cerebellum however, Leiner et al. (1986) proposes that parallel evolutionary enlargements of the prefrontal cortex and the cerebellum, simultaneously with changes in connectivity between them, provide a neural basis for the cerebellum's potential contribution to higher level cognition. Moreover, recent theoretical accounts posit that cerebellar cortical circuits store forward-models of cerebral cortical information processing that facilitate the automated execution of those processes, whether in motor *or* cognitive domains (Ramnani, 2006, 2014). The modular anatomical organisation of the cortico-cerebellar system suggests that these forward models in both motor and cognitive domains must be located in distinct areas of the cerebellar cortex. It has long been thought that the cerebellar cortex receives inputs from the cerebral cortex via the pontine nuclei and returns connections via the cerebellar nuclei and the thalamus to the same areas (Dum & Strick, 2003; Middleton & Strick, 2001). Studies using both anterograde and retrograde cross-synaptic neurotropic tracer technology have revealed the details of connectivity between areas of the frontal lobe and specific locations in the cerebellar cortex.

In non-human primates, cerebellar cortical lobules HIV-HVI, HVIIB and HVIII receive inputs from the motor cortex via the pontine nuclei and send their outputs back to the primary motor cortex (area 4) via the cerebellar nuclei and the thalamus. Targets of the prefrontal cortex (Walker's Area 46 and Petrides and Pandya's area 9/46) include lobule HVIIA (whose components are Crus I and Crus II), as well as vermal parts of lobules VII, IX and ventral parts of the dentate nucleus (Middleton & Strick, 2001; Kelly & Strick, 2003). These connections suggest that 'prefrontal' and 'motor' cortico-cerebellar loops form independent, closed-loop, circuits that operate in parallel. This gives rise to the possibility that separate forward

models in these different cerebellar cortical areas might operate independently (Ramnani, 2006).

The hierarchical organisation of the frontal lobe cortex has been discussed above, in which information cascades through prefrontal, premotor and primary motor areas such that the abstract goals of actions are elaborated into the movements. The connections of area 46 and the primary motor cortex with the cerebellar cortex are well known but the connectivity of the premotor cortex with the cerebellar cortex is less well understood. Ramnani (2012) suggested that since the cortico-cerebellar system is topographically organised, and the premotor cortex is anatomically intermediate to prefrontal and primary motor areas, the premotor cortex might be connected with the superior portions of Crus I since this is intermediate to cerebellar targets of prefrontal and primary motor areas. It is suggested, then, that processing of information related to action (at a higher level of abstraction, e.g., rules) may be used to instruct and guide actions.

It is essential to note that few reports have been made regarding the mapping of the premotor cortex and its connections to the cerebellar cortex but it is highly probable that the same cerebellar lobules project to both the primary and premotor cortices. The premotor cortex lies in between the prefrontal and primary motor cortices. Connections of the prefrontal and primary motor cortices with the cerebellar dentate appear to be topographic as dentate neurons connected to the premotor cortex are spatially interposed between those connected to the prefrontal and primary motor cortices (Orioli & Strick, 1989) it may be possible therefore, that connections from the cerebral cortex to the cerebellar cortex may be topographic in a similar fashion. Similarly to Kelly and Strick (2003), Lu et al., (2007) reported connections of cerebellar lobules HIV-HVI after injecting retrograde trans-synaptic tracer into areas

of the precentral cortex. They did not however, report how they identified the boundary between areas 4 and 6. It may be speculated therefore, that tracer was injected into both the pre- and primary motor cortices. Furthermore, Coffman et al., (2011) reported that the primary and medial premotor cortex both projected to vermal lobules VB-VIIIB. These results suggest that there is a considerable degree of overlap in primary motor and premotor projecting lobules.

In humans, resting state functional-connectivity demonstrates that lobule HVIIA (Crus I and Crus II) contains similar types of maps for prefrontal and posterior-parietal cortices and was termed the *supramodal zone* (O'Reilly et al., 2010) whilst lobules HV, HVI and HVIII contained overlapping functional connectivity maps with the somatosensory, visual and auditory cortices and was therefore termed the *primary sensorimotor zone* (O'Reilly et al., 2010). Other research using functional connectivity MRI (fcMRI) in humans demonstrated seed correlations between the motor cortex, lobules HV and HVIIIB, as well as the left dorsolateral prefrontal cortex (DLPFC, which is referred to as area 46 in the present work) with Crus I and Crus II, right DLPFC with Crus I, Crus II and lobule VIIB, the left anterior prefrontal cortex (APFC) with lobule VI and Crus I, the right APFC with lobules VI, VIIB and Crus II, and finally medial prefrontal cortex with Crus I (Krienen and Buckner, 2009).

The connections of the cerebellum to both the prefrontal and motor indicate that the cerebellum could support both the automation of cognitive and motor behaviours via the acquisition and storage of forward models. Indeed, the results of research assessing the contribution of the cerebellum in higher-order cognition are consistent with the connections of the prefrontal cortex (area 9/46) to Crus I and Crus II. Balsters and Ramnani (2011) conducted an event-related fMRI study using first-order rule learning paradigms. A first-order rule would specify a relationship between

a stimulus and the required action. The authors measured activity time-locked specifically to processing a rule (symbolic instruction cues) and subsequent action (movement of individual digits in a button press task). Results showed significant activation related to the processing of symbolic instructive cues in cortical prefrontal area 9/46 and corresponding cerebellar lobule HVIIA, specifically Crus I. Activity in Crus I decreased more rapidly for rules that were learned via non-ambiguous feedback (a cue that indicated whether their response was correct or incorrect) in comparison to ambiguous. The learning related decreases were consistent with characteristics of Hebbian learning mechanisms that underpin the feed-forward control of behaviour (Ramnani, 2006; Balsters & Ramnani, 2011). To investigate the processing of information at a higher level of abstraction, Balsters et al. (2013) implemented a second-order rule learning task where learning a set of primary rules specifies a set of secondary rules, subsequently dictating action. Activity in Crus I and Crus II evoked by second-order rule learning suggest that the cerebellum is a candidate area for the automation of cognitive operations that guide motor execution through acquisition and storage of forward models (Balsters & Ramnani, 2008; 2011; Balsters et al., 2013). These results are particularly telling as they present evidence that motor behaviours are rule guided and in-turn inform perceptual-motor hypotheses.

Cerebellar involvement in verbal short-term memory in healthy participants has been investigated extensively (Chen & Desmond 2005a,b; Chein & Fiez, 2010; Desmond et al., 1997; Durisko & Fiez, 2010; Kirschen et al., 2005, 2010; Marvel & Desmond, 2010a; Hayter et al., 2007; Ng et al., 2016; Peterburs et al., 2016, 2019; Tomlinson et al., 2014). The majority of these studies were closely aligned with the phonological loop model, attributing activation in the inferior cerebellum (lobules HVIIB and HVIIIA) to reciprocal connections with a phonological store in cortico-

temporo-parietal regions (Chen & Desmond, 2005a,b; Desmond et al., 1997; Kirschen et al., 2005, 2008, 2010; Ng et al., 2016). However, much like the studies reviewed in Section 1.5, activation of both lobules HVIIB and HVIIIA proposed to be involved phonological storage was not consistently observed throughout encoding, maintenance, and retrieval phases of a Sternberg task (Sternberg, 1966), which is at odds with the predictions of the phonological loop model. In contrast, interpretations of cerebellar involvement (specifically, lobules HVI and HVIIA) in an articulatory rehearsal process based on the same studies; Chen & Desmond, 2005a,b; Desmond et al., 1997) are in line with evidence of anatomical connections between the cerebellum and cortical frontal regions as well as cerebellar involvement in temporal and sequential aspects of motor behaviours (Leggio & Molinari, 2015).

The view that the cerebellum was supplied with phonological information supporting a phonological store was based on a short-term memory study using visual-verbal stimuli: Desmond et al. (1997) required subjects to remember six (high load) or one (low load) letters across a brief delay, or to covertly 'rehearse' letters (rehearsal control condition, also contrasting high vs. low load) as they were presented to them at a similar rate to rehearsal in the memory condition. After the delay in the memory condition, a singular probe stimulus was shown which matched either one of the letters in the high load condition, the letter in the low load condition, or was not a match. Subjects were instructed to provide a manual response when a match occurred. The rehearsal condition presented items-to-be-rehearsed four times and included no delay. Whilst activation in lobules HVI, HVIIA, and HVIII were observed in the results of the high vs. low memory contrast and the rehearsal control condition, activation of lobule HVIIB was only observed in the results of the memory contrast. This was interpreted as reflecting a load effect that taxed phonological storage. A loop involving

temporo-parietal regions and cerebellar lobule HVIIB was proposed to be used for error correction of items said to be held within a phonological short-term store. And yet the study did not in fact assess activation across the cortex.

Using the same experimental design, Chen and Desmond (2005a,b) aimed to extend those findings and did on this occasion assess activation across both the cortex and cerebellum. Activation in right lobule HVIIB was again interpreted as supplying a phonological store with information as the lobule was co-activated with bilateral BA 40 in inferior parietal lobe (Chen & Desmond, 2005a). However, the focus of the theorising remained on the left hemisphere as the location of the store. The activation of lobule HVIIB was suggested to "influence the motor trajectory of the phonological loop based on the internal guidance of the phonological store" (Chen & Desmond, 2005a, p. 337) by relaying information to the parietal lobe (Clower et al., 2001). Chen and Desmond (2005a) further supported the claims of the phonological store connections with lobule HVIIB through reference to previous research that had suggested that regions of the parietal cortex (rather than frontal cortical regions) were more likely to influence the inferior cerebellum (Brodal, 1979; Desmond, 2001; Schmahmann, 1996; Schmahmann & Pandya, 1997a,b).

Critically, the interpretation of regions involved in phonological storage influencing the trajectory of an articulatory rehearsal process appear to be at odds with the cognitive-psychological characterisation of the phonological store. The phonological loop model does not state that the store guides the articulatory rehearsal process but rather that the active articulatory process feeds the phonological store (in the case of visual-verbal input) and refreshes item representations within the store (regardless of modality). The authors themselves note that their interpretations should be taken with a degree of caution because differences between the memory and

rehearsal control task may have resulted in activation of lobule HVIIB specific to the memory condition: it was highly likely that the repeated stimulus presentation during the 'rehearsal' control task resulted in habituation that may not have required error related feedback from lobules HVIIB that were observed across the temporal delay in the memory condition.

A further study by Chen and Desmond (2005b) attempted to isolate activity specific to each task phase in the Sternberg task. Activation of lobule HVIIB and HVIIIA was observed during both presentation and retention phases but not retrieval. This in itself is problematic with regard to phonological store theory as regions contributing to phonological storage should be observed during all three task phases. In comparison, lobule HVIIB was only observed to be active during the presentation phase of a similar Sternberg task (Peterburs et al., 2016) and a verbal delayed serial recall task (Durisko & Fiez, 2010). Further inconsistencies were also observed in the activation across the inferior parietal lobe depending on task phase: BA 40 was only observed during the maintenance phase (Chen and Desmond, 2005b) or presentation (right inferior parietal lobe) and retrieval (left supramarginal) (Marvel & Desmond 2010a) during other verbal short-term memory tasks. Again, such results are at odds with the phonological loop model: If activity of regions across the inferior parietal lobe, or in cerebellar lobule HVIIB were involved in phonological storage, then their activation should be observed throughout the presentation, maintenance and recall of verbal items (cf. Chein & Fiez, 2001).

In contrast, activation of lobule HVI has been found to be persistent throughout all phases of five different tasks: delayed serial recall, covert speech, overt speech, covert tapping and overt tapping **(**Durisko & Fiez, 2010) and suggests that lobule HVI is involved in the planning and execution of sequential motoric behaviour. Indeed,

there is continued support for the involvement of lobule HVI and Crus I in articulatory processes outside the domain of short-term memory tasks (Murdoch, 2010; Peeva et al., 2010 Stoodley & Schmahmann, 2009, 2010). More specifically, lobule HVI has been associated with lip and tongue movements, suggesting that the lobule may contain forward models of the vocal tract and mouth for articulation (Callan et al., 2007). Moreover, results of cerebellar TMS studies suggested that interference to lobules HVI and Crus I impact an articulatory trajectory in verbal Sternberg tasks (1966). Single-pulse TMS administered to lobule HVI and Crus I after letter presentation showed no effects on accuracy but reaction times on correct trials were significantly increased (Desmond et al., 2005). In contrast, continuous theta burst stimulation (cTBS) administered to the same lobules following stimulus presentation attenuated performance (Tomlinson et al., 2014). Taking the results of these studies (Desmond et al., 2005; Tomlinson et al., 2014) together with evidence of anatomical connections of lobule HVI and Crus I with the motor and prefrontal cortices respectively (Kelly & Strick 2003; Lu et al., 2007), it is likely that TMS reduces predictive control of the articulatory trajectory necessary for successful performance in verbal short-term memory tasks. In line with the earlier discussion of Crus I contribution to higher-order cognition (Balsters & Ramnani, 2011; Balsters et al., 2013), it should be considered that consistent co-activation of Crus I with lobule HVI may indicate the operation of forward models predicting the responses of premotor cortex to higher-level commands issued by the prefrontal cortex (Ramnani, 2006).

It would appear, then, that the assessment of cerebellar involvement in verbal short-term memory may have been overly influenced by the dominant phonological loop model (Baddeley & Hitch, 1974). Claims that lobule HVIIIA supports the maintenance of representations within a store appear to be incompatible with

phonological store theory. Instead, suggestions that lobule HVI and Crus I contain forward models for an articulatory rehearsal process are more consistent with the focus on motor planning in the perceptual-motor account of verbal serial short-term memory.

1.7 The Present Empirical Work and Hypotheses

Recent behavioural research has demonstrated that performance in verbal serial short-term memory tasks can be explained in terms of articulatory planning and perceptual organisation processes without the need to posit a specialised phonological store (the *perceptual-motor account*; Hughes & Marsh, 2017; Hughes et al., 2009, 2016; Jones et al., 2006, 2004; Macken et al., 2016; Sjöblom & Hughes, 2020). The empirical work reported in the present thesis (Chapters 3 and 4) examined the cognitive neuroscience of verbal serial short-term memory and learning specifically from the standpoint of the perceptual-motor account for the first time.

The present empirical work comprised two fMRI verbal serial recall experiments, each with two sub-experiments (A & B). While the method is described in detail in Chapter 2, the central features of the experiments are also described here. In Experiment 1A, participants were auditorily presented with a short sequence of letters presented at the rate of 700 ms per item. Following sequence presentation and a short temporal delay, subjects were required to reconstruct the order of a fragment of the just-presented sequence using a four-button response-box (for details see Chapter 2, Section 2.2.3). The lists in this sub-experiment were either phonologically similar to one another (‘B’ ‘C’ ‘D’ ‘T’ ‘P’ ‘V’ ‘G’) or phonologically dissimilar (‘F’ ‘K’ ‘L’ ‘R’ ‘Y’ ‘H’ ‘Q’). A phonological similarity effect—poorer serial recall of the similar items—would, from the perspective of the perceptual-motor account, provide independent behavioural corroboration for the assumption that the task promoted an

articulatory planning/rehearsal strategy (e.g., Jones et al., 2004). And from the perspective of the phonological loop model, a replication of the phonological similarity effect would be taken as evidence that the passive phonological store was utilised (e.g., Baddeley & Larsen, 2007).

A further critical feature of Experiment 1A was that in some trials (regardless of the phonological similarity manipulation), subjects were presented with a 'Go' cue following sequence presentation and before the temporal delay and this meant that they would be required to recall (a fragment of) the sequence following the temporal delay. For other trials, a 'No-Go' cue was presented at the same point, which informed subjects that they would *not* be required to recall the sequence following the temporal delay. It was assumed, therefore, that subjects would continue engaging in articulatory rehearsal during the temporal delay in the 'Go' trials—in preparation for recall—but not during the temporal delay in the 'No-Go' trials.

It is hypothesised that the mid-portion of the middle frontal gyrus (putatively area 46) and cerebellar cortical Crus I and Crus II of lobule HVIIA (Kelly & Strick, 2003; Krienen & Buckner, 2009; O'Reilly et al., 2010) will be activated during the temporal delay period of 'Go' trials compared to the temporal delay period of 'No-Go' trials. This would be taken to reflect the maintenance of motor-related information in abstract terms (higher-level representations of the goals of the action; Passingham, 1996) as subjects engage in articulatory rehearsal in support of the imminent recall prompt. Activation of the premotor cortex (BA 6) and its cerebellar targets—lobules HIV-HVI, HVIIB and HVIIIA (Kelly & Strick, 2003; Krienen & Buckner, 2009; Lu et al., 2007—is also predicted when comparing haemodynamic activation during the temporal delay from 'Go' and 'No-Go' trials and would be interpreted to reflect the execution of specific motor commands during articulatory rehearsal. As previous

studies have reported that activation of the inferior frontal gyrus (BA 44; pars opercularis) reflects the generation of articulatory codes and the mapping of such codes to orofacial musculature (Hickok & Poeppel, 2004; Papoutsi et al., 2009; Price, 2012; Rauschecker & Scott, 2009) activation of this region is also expected in the results of the same contrast.

As the perceptual-motor approach proffers that the assembly of an articulatory plan begins as soon as the items begin to be presented (e.g., Hughes et al., 2009; Macken et al., 2016), a conjunction analysis is predicted to demonstrate that regions underpinning motor planning (across cerebral and cerebellar cortices as in the previous hypothesis) will consistently be active during both presentation of sequences and across the temporal delay (of Go trials) prior to recall. Critically, no region in the parietal lobe—that could ostensibly be the location of a phonological store—is expected to be consistently active across both task phases (see also below).

Activity during presentation will also be contrasted with that during the temporal delay (of Go trials). Regions associated with auditory-perceptual or auditory-motor mapping processes are predicted to be active, over and above those observed in the conjunction analysis. These perceptual regions include the auditory cortex and planum temporale. If such activation is observed alongside other regions that can be explained by recourse to motor planning or perceptual-motor mapping—whilst no region consistent with the functional and supposed anatomical characteristics of a phonological store is found to be active—the results would support the contention that verbal serial short-term memory is parasitic on motor and perceptual processes and that there is no distinct phonological store.

The second main aim of the current empirical work was to assess the neural basis of long-term verbal sequence learning. Such learning has been described as the

evolved function of the phonological store (Baddeley et al., 1998). However, the current approach suggests that long-term learning may be better understood in terms of (haemodynamic changes in) brain regions associated with motor planning (cf. Sjöblom & Hughes, 2020). Experiment 1B—which followed on immediately after Experiment 1A using the same sample of subjects—therefore employed the Hebb sequence learning paradigm in which a repeating sequence is interspersed amongst non-repeating 'filler' sequences (Hebb, 1961). Recall of the repeating Hebb sequence typically increases in accuracy across the course of the experiment whilst recall of non-repeating filler sequences remains unchanged. This paradigm was used to assess whether the neural signatures of articulatory rehearsal as revealed in Experiment 1A, changed over time as a function of repetition learning. It was predicted that long-term learning of a repeating sequence (Hebb sequence learning; Hebb, 1961) would be indicated by increased behavioural accuracy concomitantly with an excitability decrease (decrease in BOLD signal activity, Albus, 1971) in cerebellar lobules HIV-HVI, HVIIB and HVII. Such results would indicate that the articulatory plan associated with the repeating sequence becomes increasingly automated and suggests that forward models of cortical premotor information processing held in the cerebellum facilitate the automatic execution of covert articulatory processes (Ramnani, 2006). An effective approach, extensively used to investigate the acquisition of motor skills, is to use parametric methods (details of which are discussed in Chapter 2, Section 2.7.3) (Ramnani et al., 2000; Sakai et al., 2002b; Toni et al., 2001). Such methods allow for statistical manipulation of a given variable and thus allow assessment of the extent to which the amount of activity changes across a time course and is used in the present empirical work to assess changes associated with long-term verbal sequence learning.

Experiments 2A and 2B were identical to Experiments 1A and 1B, respectively, but the to-be-remembered items were presented *visually*. Conducting the same experiments using visual-verbal stimuli was designed to further adjudicate between the phonological store theory and the perceptual-motor account: On the former theory, there should be a single region that is consistently active not only across task phases (as in Experiment 1) but also across presentation modalities (auditory – Experiment 1, and visual – Experiment 2) and the activity in this region should not be easily attributable to the contribution to performance of non-storage processes such as motor, modality-specific perceptual, or language processes. The current phonological-store based view is that the store is localised to the inferior parietal lobe (BA 40) and therefore, from this perspective, it is this region that is most likely to be found to be active across both task-phases and presentation modalities in the current experiments (Baddeley, 2003; Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996). But with the advent of increasingly more refined neuroimaging methods and connectivity analyses, it is becoming clear that the parietal cortex is a computationally dense region, connected to numerous other areas of the neocortex as well as to the cerebellum and that it could, therefore, support a whole host of functions (Richter et al., 2019) in the context of verbal serial short-term memory performance. In contrast, the perceptual-motor account predicts that the only regions that will be consistently activated across task-phases with visual presentation will be the same motor planning systems predicted to be active across task-phases in Experiment 1 with auditory sequences. But in addition, key differences in activations are of course predicted for the presentation phase as a function of presentation modality, reflecting the different brain systems involved in visual and visual-motor mapping processes (Experiment 2) compared with auditory and auditory-motor mapping processes

(Experiment 1). Further details regarding the hypotheses relating to the role of modality-specific (visual) perceptual and perceptual-motor mapping are postponed until the relevant chapter (Chapter 4). Finally, to the extent that motor planning is argued on the perceptual-motor account to support both auditory-verbal and visual-verbal sequence learning (Sjoblom & Hughes, 2020), it is predicted that the same decrease in cerebellar lobule activity as a function of Hebb repetition as predicted for Experiment 1B (auditory) will be evident also in Experiment 2B (visual).

CHAPTER II: GENERAL METHODOLOGY

2.1 Introduction

This chapter describes the general methodological approaches used across the two experiments reported in this thesis, each of which comprised two fMRI experiments (Experiments 1A and 1B, and Experiments 2A and 2B). The construction of an experimental design with an appropriate trial structure was necessary to investigate neural activation across a temporal delay during a verbal serial recall task. The experiments focus on two complimentary roles of the same articulatory planning process. The first is the role of such planning in the explicit short-term retention and reproduction of a verbal sequence and the second is its role in the longer-term incidental learning of a repeating verbal sequence. It was therefore essential to design a within-trial structure that was identical across the A and B components of each experiment so as to probe the same motoric process. In this way, if dynamic sequence-learning related changes (cf. Experiments 1B and 2B) are observed in the same regions found to be active during articulatory rehearsal engaged for short-term recall (cf. Experiments 1A and 2A), this would support the view that the same articulatory planning process supports long-term sequence learning. A further requirement of the experiment design was to ensure suitability for both auditory and visual modes of sequence presentation that would also enable casual comparison across experiments regarding perceptual and motor processing.

Functional magnetic resonance imaging (fMRI) is deemed an appropriate methodology for the current investigations, particularly due to its high spatial resolution. The ability to isolate blood oxygenation level dependent (BOLD) signal responses within specific neuroanatomical regions will enable a comparison of results

from the present investigations to those claiming to have localised a phonological store in a region functionally and anatomically separate to any motor or perceptual regions (see Chapter 1, Section 1.5). FMRI also enables the ability to assess haemodynamic changes across extended periods of time which is particularly suitable for testing the long-term learning hypotheses of the present investigations. Moreover, the logistics of conducting fMRI experiments are more efficient than those of other methods outlined below, meaning that a relatively high volume of subjects can participate within a given day.

Whilst the majority of early neuroimaging studies used PET, such an invasive method was not deemed appropriate for pragmatic reasons in the current context. Other techniques such as transcranial magnetic stimulation (TMS) and electric encephalography (EEG) were also inappropriate for the current investigations. TMS can temporarily stimulate or interfere with neural processing of a specific area but such a method would not differentiate between processes (and regions) hypothesised to be involved in verbal short-term memory from the perceptual-motor perspective nor afford an assessment of previous attempts to localise a phonological store. This is because TMS does not present the opportunity to visually assess the concomitant activation of regions across the brain that could indicate the type(s) of processing and systems involved. Additionally, the relatively poor spatial resolution of EEG would not enable the testing of the relatively anatomically-specific hypotheses of interest in the present thesis.

Using fMRI (a non-invasive neuroimaging technique) enables the identification of changes in cerebral blood flow (CBF) throughout the brain that are indicative of changes in neural activity. The suggestion that changes in CBF are linked to changes in neural activity emerged towards the end of the 19th century (Roy &

Sherrington, 1890). It was not until the start of the 20th century, however, that BOLD fMRI was developed as a method of investigating and observing changes in the brain localised to a particular region. These observable changes occur as a result of neuronal firing, a metabolically demanding process requiring oxygen and glucose. Blood vessels in close proximity to firing neurons then begin to regulate CBF, thus increasing blood flow to particular regions where neural firing has occurred. For example, when subjects are required to visually fixate on a black and white checkerboard, increased CBF is observed in specific regions of the brain (such as the primary visual cortex; Kim et al., 1999). FMRI is able to detect any such changes in blood flow due to changes in magnetic susceptibility that arise from the consumption of oxygen. Blood without oxygen (deoxyhaemoglobin) is more paramagnetic than blood with oxygen (oxyhaemoglobin). Oxygen extraction and metabolisation are accompanied by a further increase in CBF but, critically, the increase in CBF is larger than the amount of oxygen extracted (Buxton & Frank, 1997; Buxton et al., 2004). This results in an increase of the ratio between oxy- to deoxyhaemoglobin in areas where neuronal firing has occurred and change in magnetic susceptibility is then detectable using fMRI. Thus, it is important to note that fMRI measures neural activity indirectly via the BOLD signal rather than neuronal firing itself (Logothetis, 2008).

Several methodological steps were taken to ensure that the experimental design and chosen neuroimaging technique were optimal for the present studies. First, estimated timings based on the proposed experimental design were used to generate design matrices in SPM12. Second—having confirmed that the design was statistically feasible for General Linear Modelling (GLM) fMRI statistics (see section 2.7.3)—pilot experiments were conducted in a simulated MRI environment to assess the viability of the behavioural paradigms, especially as short-term serial recall and

long-term learning would be assessed via a unique order reconstruction task (see details in Section 2.2.3). It was necessary to evaluate the overall design of the experiment as well as the precision of recorded timings for stimulus presentation and subject responses. The pilot experiments also allowed testing of experiment scripts and hardware, where manual responses and calculation of recall accuracy were reviewed. Following pilot data collection and analysis, the optimal design was then used for pilot fMRI data acquisition. Once the pilot fMRI data had been examined, both the experiment design and GLM were considered feasible and full dataset acquisition commenced.

2.2 General Design and Procedure

Two fMRI experiments were conducted: Experiment 1 involved presenting to-be-remembered sequences auditorily while Experiment 2 involved presenting them visually.

2.2.1 Experiments 1A and 2A

Sub-experiment A of each experiment consisted of 48 trials. As shown in Figure 4A, during each trial, a sequence of seven letters was presented (auditorily in Experiment 1 and visually in Experiment 2). Half of the sequences (n = 24) consisted of random orderings of the same set of seven phonologically similar letters (‘B’ ‘C’ ‘D’ ‘T’ ‘P’ ‘V’ ‘G’) and the other half (n = 24) consisted of random orderings of the same set of phonologically dissimilar letters (‘F’ ‘K’ ‘L’ ‘R’ ‘Y’ ‘H’ ‘Q’). The order of the letters for the 48 sequences was randomised anew for each subject. For each list-type (Phonologically similar, Phonologically dissimilar), half the trials (n = 12) were ‘Go’ trials in which, following the last to-be-remembered item, a ‘Go’ signal appeared to inform subjects that, following a short temporal delay, they were to recall the list. In

the other half of trials (n = 12), the signal read ‘No-Go’, which informed subjects that they would not need to recall the sequence (Figure 4B). ‘No-Go’ trials nevertheless had the same temporal delay as the ‘Go’ trials, though of course subjects were not required to retain a sequence for recall in this case but instead were instructed merely to wait until a fixation cross indicated the imminent start of the next trial (note that no deceptive trials were included where subjects were cued with ‘No-Go’ but then required to recall the sequence). This resulted in four different trial-types, with 12 trials of each: Phonologically similar—Temporal delay (Go), Phonologically similar—Temporal delay (No-Go), Phonologically dissimilar—Temporal delay (Go) and Phonologically dissimilar—Temporal delay (No-Go). For each subject, these four trial-types were mixed in a 48-trial block in a pseudo-randomised fashion with the constraint that no trial-type was presented more than three times in succession.

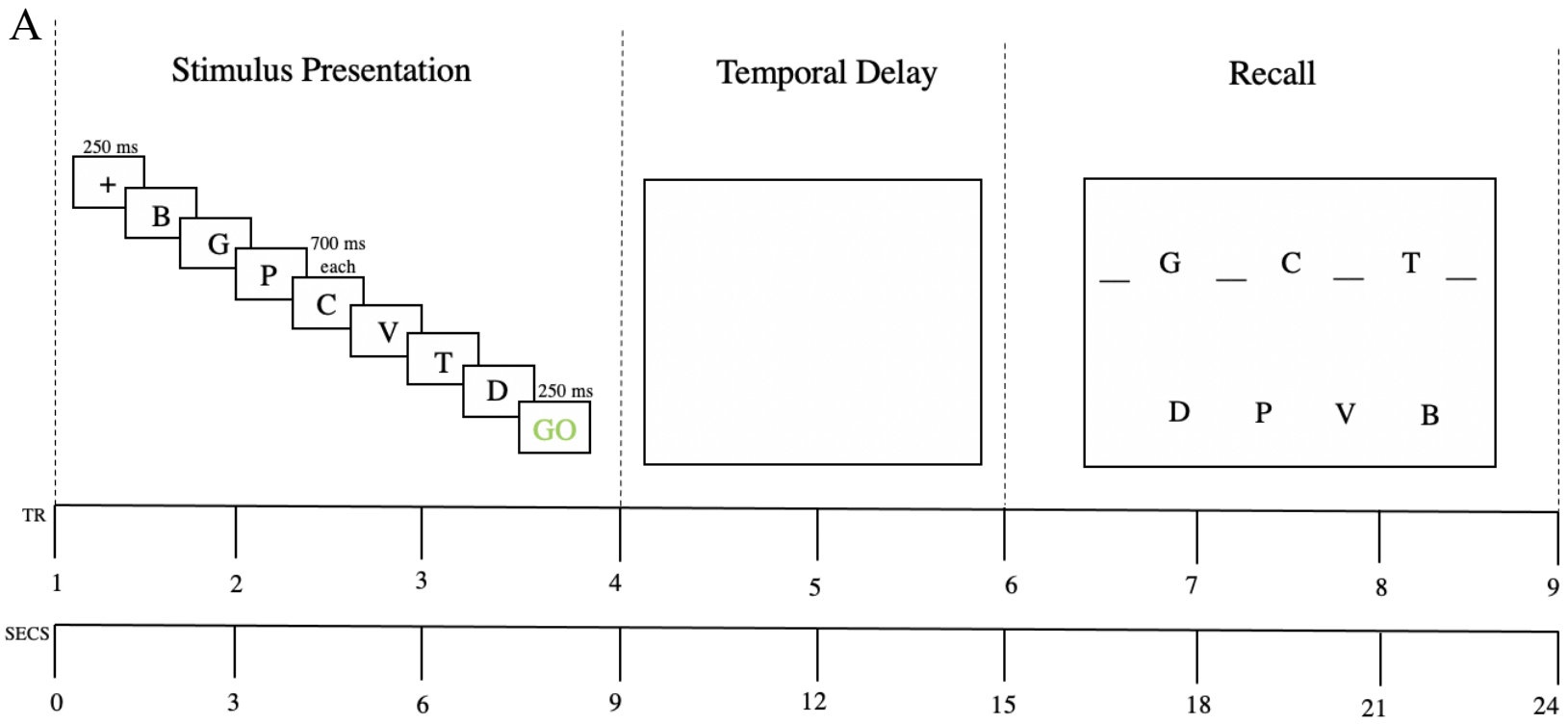
2.2.2 Experiments 1B and 2B

Experiment B always succeeded Experiment A with no break in between. Experiment B consisted of 60 trials, all of which were ‘Go’ trials. The trials comprised 43 non-repeating ‘Filler’ sequences and 17 instances of the same (i.e., repeating) ‘Hebb’ sequence. All sequences in this experiment consisted of the seven phonologically similar letters used in Experiment A. The first pilot experiment (using visual stimuli; Appendix A, Section 6.4.2.1) had shown that using phonologically dissimilar sequences resulted in a relatively weak Hebb effect due to recall level being relatively high from the outset, that is, before any Hebb repetition. Phonologically similar sequences were used, therefore, to reduce initial performance level so as to increase the chances of observing an enhanced level of recall due to Hebb repetition. The order of the letters in a sequence was randomised anew for each Filler trial but, necessarily, remained the same for each Hebb sequence. A different Hebb sequence was generated

for each subject and no two Filler sequences were the same for any subject. In an attempt to reduce expectancy of the recurring Hebb sequence, trials were spaced either two or three Filler sequences apart. Whilst the Hebb sequence learning paradigm traditionally involves consistent spacing of the Hebb sequence, in the context of an fMRI experiment it was necessary to reduce any expectancy of such trials. This is because the BOLD-signal can be influenced by subjects' expectation of an upcoming (and repetitive) stimulus such as Hebb sequence presentation, rather than processes related to the learning of the sequence per se. The Filler-Hebb trial manipulation was pseudo-randomly generated when designing the experiment and then fixed across all subjects.

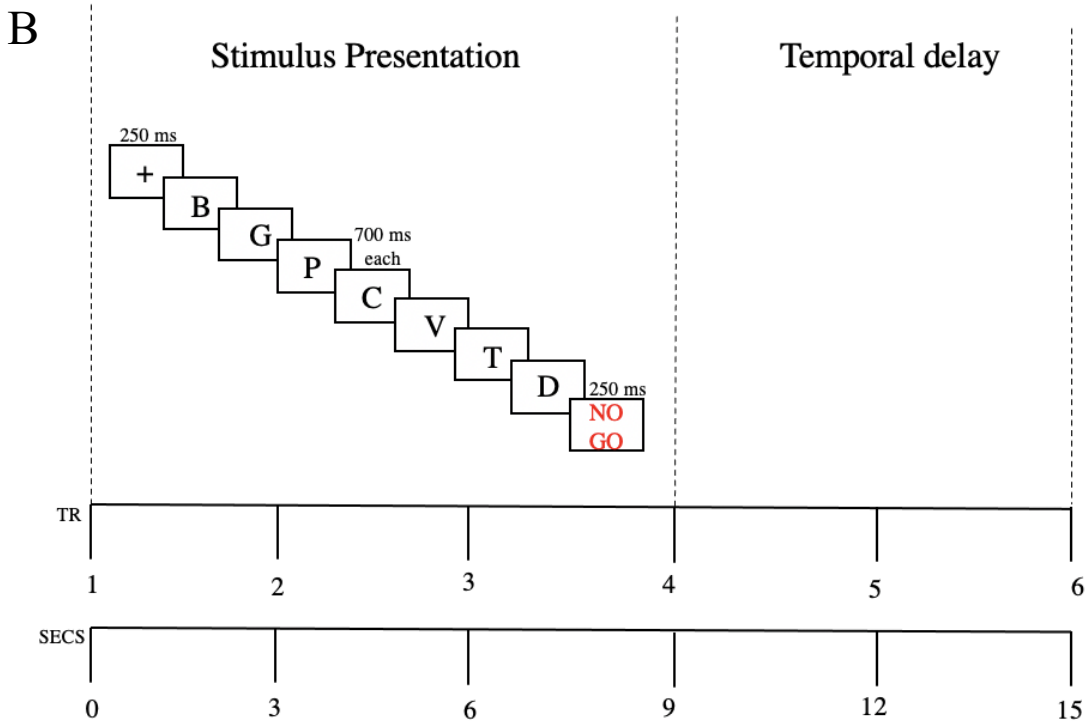
Figure 4

Panel A: Structure of a 'Go' trial. Panel B: Structure of a No-Go trial.



Note. The dashed lines delineate the repetition times (TRs) of the MR pulse between which the onset and offset of an event (stimulus presentation; temporal delay; recall) could occur. The dashed lines further outline the duration of an entire event in seconds and corresponding number of TRs. The duration of the various stimuli during stimulus presentation is shown above the stimuli; of particular importance, the to-be-remembered letters were presented one at a time for 700 ms each. The signal 'Go' in green or 'No Go' in red (Panel B) indicated whether a response would be required following the end of the temporal delay. The

square in the temporal delay phase represents a blank screen. Following the temporal delay, three of the initially presented letters were re-displayed in their original serial positions. The remaining four were re-presented underneath in a random order. Subjects were required to reconstruct the correct order of the remaining letters in the sequence.



Note. Only relevant only for Experiments 1A and 2A. The blank Temporal delay screen was followed by a fixation cross marking the start of the next trial.

2.2.3 Within-trial structure

As shown in Figure 4A, event 1 ('Stimulus Presentation') was identical across 'Go' and 'No Go' trials other than the signal following the last to-be-remembered item. The start of each trial was indicated by a black fixation cross shown for 250 ms in the centre of the screen. In Experiment 1, the to-be-remembered items were presented via earphones and the screen remained blank during stimulus presentation. Each of the seven letters was presented one at a time for a duration of 500 ms with an inter-stimulus interval of 200 ms. In Experiment 2, the to-be-remembered items were visually presented, one at a time, for a duration of 700 ms each (with no inter-stimulus interval). Thus, the overall rate of letter presentation was the same across the two experiments (one every 700 ms). The order of the letters for each trial other than Hebb trials was randomised. Following sequence presentation, subjects were shown a signal for 250 ms: either 'Go' in green, or 'No-Go' in red. This indicated to the subject whether or not they would be required, in a few seconds, to attempt to recall the sequence, or if they were to merely wait until the fixation cross marking the start of the next trial. In the recall phase, three of the seven letters were simultaneously presented in their original serial positions. Underneath the sequence fragment, the four remaining letters from the just-presented list were simultaneously presented in a random order. Subjects were required to select the four letters in the order in which they had occurred within the just-presented sequence using a four-button response box. For instance, in the example given in Figure 4A to be correct, the subject would need to first press the rightmost button to fill in the first blank space with the 'B', followed by the second from left button to fill the next blank with the 'P', and so on. It is important to note that subjects were required to retain all seven items in the sequence as they were unaware of which four blank spaces, they would need to fill on any given trial until

the onset of the recall phase (though the first serial position always needed to be filled; see Appendix A, Section 6.2.1). The method of sequence-fragment completion as opposed to full serial recall was a novel way of examining serial recall behaviour, necessitated by the fact that the response box used to collect subjects' responses had only four response-buttons. This method turned out to be particularly suitable for present purposes, however, as responses could be made with just the one hand—thereby limiting the amount of bodily movement in the scanner—and it meant that the mapping of stimuli to buttons was relatively simple.

2.2.4 Temporal Jittering and Durations

Temporal jitters are variable delays included in experiment designs to vary the onset of events within a given trial as well as between trials. Separating events using temporal jittering allows the BOLD-signal associated with each event to be isolated and ensures that subjects are unable to predict the timing of event onsets, meaning BOLD activity could not be related to predictive processes. Furthermore, using jittering between events ensures that the sampling of haemodynamic responses is optimal for repeated stimuli, required processes and that the data are appropriate for GLM statistics. The method of jittering with variable delays is particularly useful for GLM statistics, as once a first-level model (see Appendix C) has been generated, the correlation coefficients between each regressor in the model can be calculated. Low correlations between regressors indicate that there is little shared variance and that events are separated enough in time so that the BOLD-signal could be attributed to processes that are uniquely implicated during a given event if there are such.

Based on the duration of stimulus presentation in event 1, the duration of the temporal delay (event 2), and the allotted time for subjects to respond (event 3), the remaining time available to include variable delays was calculated. The delays were

included at the start of each event, meaning there could be shorter or longer intervals between the end of one trial and the start of the next, as well as following the 'Go' or 'No-Go' signal and the recall-phase on each trial. Jitter durations were generated anew for each subject. The dotted lines shown in Figure 4A and 4B indicate between which TRs the onset and offset of events were jittered. A trial began at TR 1 and the offset of event 1 occurred prior to TR 4. The duration of event 1 was 5850 ms, thus allowing for a 3150 ms temporal jitter. The jitters were split between the TR marking the start of the trial and the fixation cross, onset of the first stimulus in that event, as well as the offset of the signal and the next TR marking the onset of event 2. A 1550 ms jitter was incorporated into event 2 but this had no visible effects as subjects viewed a blank screen following the offset of the instruction signal in event 1. A jitter of 1800 ms was split between the start of event 3 and the onset of the response screen, as well as the end of the response phase and the TR marking the start of the next trial. When modelling data from the temporal delay the regressor of event 2 was extended to include the jitter following the offset of the instruction signal and prior to the onset of the recall screen. This maximised the sampling of what was hypothesised to be a rehearsal process in the absence of any stimuli. The maximum length for a 'Go' trial was 24 s and the minimum was 18 s. The length of a 'No-go' trial was always 15 s. The maximum duration of completing all 108 trials (i.e., across both Experiments A and B) was 39.6 min.

2.3 Apparatus for Behavioural Pilot Experiments

The aims of the behavioural pilot experiments were to test the feasibility of experimental designs that were thought to be statistically optimal, under conditions physically comparable to those of the 3T MRI system in the MRI unit. To replicate the MRI environment as closely as possible for the pilot experiments, a hardware set-up

was constructed including a mock scanner and MRI-relevant software were used (see Figure 5). This allowed for assessment of the within-trial structure, accuracy of the timings of stimulus display, and the responses of button boxes to be used in the scanner. The mock MRI scanner set-up consisted of the components shown in Figure 5. The dimensions of the mock scanner were comparable with those of the 3T MRI CUBIC scanner to be used for the actual fMRI experiments and included a scanner bed. The mock scanner bed had a mock head coil with a mirror attached to allow viewing of a monitor connected to the Psychtoolbox laptop. Psychtoolbox interfaces with MATLAB and utilises a set of functions for presentation of auditory and visual stimuli. Subjects therefore viewed or listened to stimuli generated with Psychtoolbox on a monitor via the overhead mirror or in-ear earphones. Visual stimuli were flipped left-to-right so that they appeared in the correct orientation. This was achieved using UltraMon software running on the Psychtoolbox laptop. Subject responses were also collected using the laptop running Psychtoolbox. Since all experimental events were synchronised to TRs, a CED1401 generated a transistor-transistor logic (TTL) pulse every 3 s to mimic TR onsets. The CED1401 also sampled the onsets, durations and offsets of all stimuli and responses via a parallel port box which were visualised on the Spike PC. The speakers played a sound file of MRI scanner noise obtained online in an attempt to replicate the actual scanner environment as closely as possible.

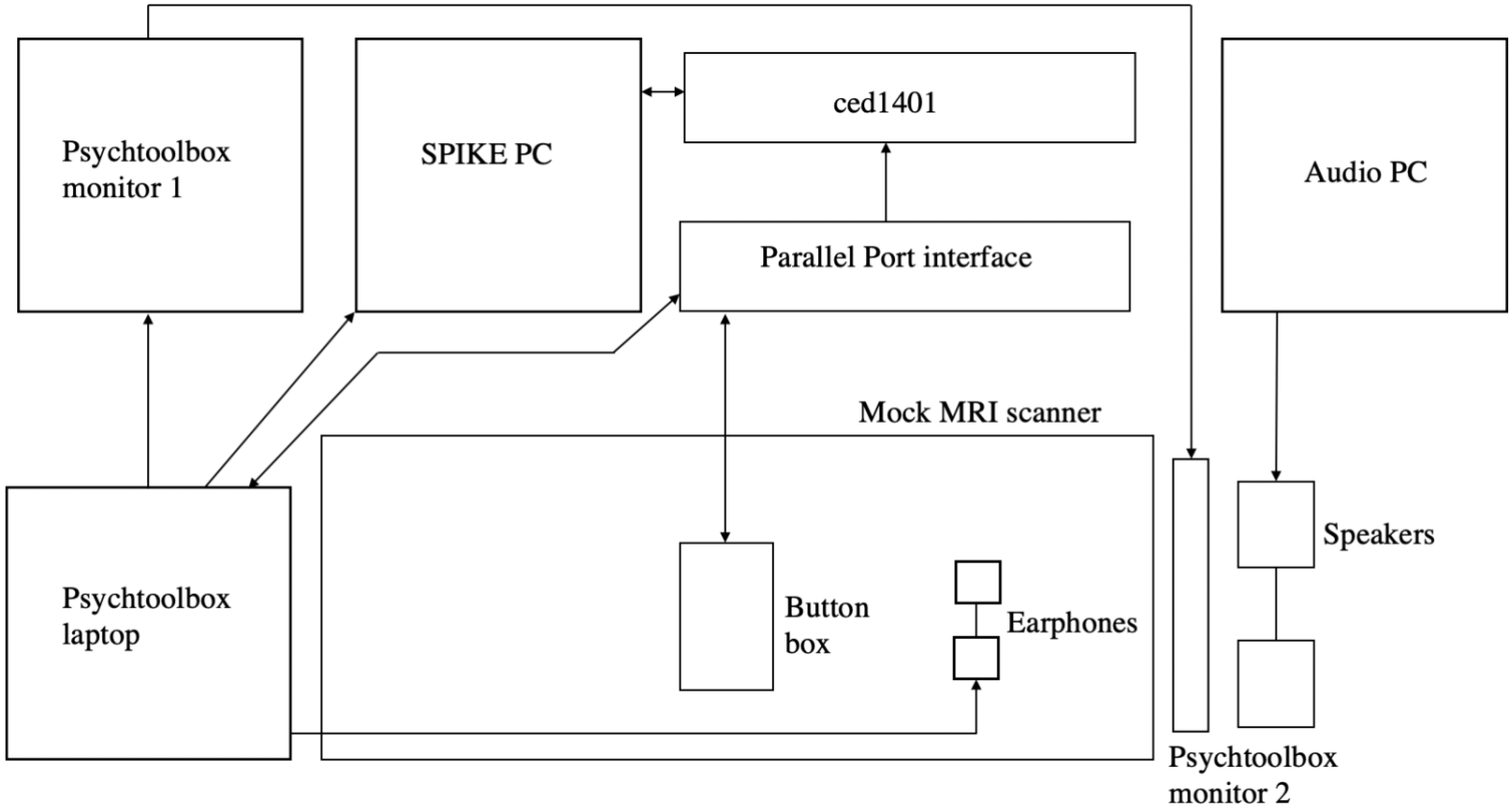
2.4 Behavioural Data Acquisition and Analysis

Prior to data acquisition, ethical approval was obtained from Royal Holloway's Ethics Committee (Code: 418-2017-28-17-05-PDJT002). Behavioural data (accuracy, experimental and response times) were acquired using MATLAB and Psychtoolbox. MATLAB scripts were used to determine which responses were correct (i.e., an item placed in the correct serial position in a sequence fragment) and order timings for

creation of individual subjects' first-level GLMs. The first script scored a response as correct only if the correct letter was placed in the correct serial position in the sequence-fragment, meaning responses were scored out of four. The second script ordered trials into separate variables based on the trial-type, in order of the onset and offset timings. These variables were then used as regressors in the first-level models for fMRI analyses in SPM12 and automatic analysis (aa). Behavioural data from the experiments were analysed using IBM SPSS Statistics 23.

Figure 5

Diagram showing the simulated MRI scanner hardware set-up for behavioural-pilot experiments.



Note. The earphones were only used in relation to the pilot experiments for Experiment 1 (auditory).

2.5 Viability for Functional MRI Experiments

To assess the viability of the experiments for fMRI, pilot experiments for all four experiments to be reported in this thesis (1A, 1B, 2A, and 2B) were conducted in the mock scanner. The results of these are reported in Appendix A. To summarise, both the phonological similarity effect (Experiment A) and the Hebb sequence learning effect (Experiment B) were replicated with both auditorily- (Experiment 1) and visually-presented items (Experiment 2), indicating that the design was fit for purpose in terms of the behavioural aspect of the experiments. Using individual subject timings from behavioural pilot data, first-level models were estimated in SPM12. From this, correlation coefficients between regressors in the first-level model were calculated. All correlations between regressors were low (within the range of 0.3 to -0.3) thus implying that there was little shared variance between experimental regressors to be used for group level analyses. To maximise the sampling of what was hypothesised to be a rehearsal process in the absence of any stimuli, the experimental design was viable for GLM statistics.

2.6 Functional MRI Experiments

2.6.1 Apparatus

Subjects lay supine in a 3 T Siemens MRI scanner and were required to wear MR-safe earphones for Experiment 1 and to wear ear plugs in Experiment 2. Subjects positioned the four fingers of their right hand on a four-button MRI-compatible response box or NATA response pad. Stimuli were back projected onto a screen behind the subject and viewed in a mirror. A dedicated scanner PC was used for stimulus presentation and behavioural data collection using Psychtoolbox in MATLAB for Experiment 1 and a separate laptop for Experiment 2. The stimulus laptop and scanner PC received TTL

pulse inputs from the MRI scanner, allowing the events of each trial to be synchronised to the onset of each scan.

A novel aspect of these experiments concerned the "Recall" event. A serial recall task typically requires subjects to output the entire sequence vocally, write out the sequence, type the sequence using a keyboard, or reconstruct the order of the whole sequence by mouse-clicking on a display showing the items re-presented simultaneously in a different order. As the current experiments were to be conducted in an MRI scanner, these traditional cognitive-psychological methods of recall would have presented considerable challenges. Whilst sparse sampling techniques are often used for the acquisition of functional data during overt vocal responses, the allotted number of scanning hours did not permit the required number of trials and subjects for suitable statistical analysis. Moreover, the technical design of a sparse-sampling experiment was outside the remit of this PhD project. Due to these considerations, it was decided that the most practical way of recording behaviour was to use a four-button, MRI-compatible response keypad. Subjects were only required to select four of the seven letters in the order in which they were presented (see Section 2.2.3). There is ample evidence that key verbal serial recall phenomena—including the phonological similarity effect (e.g., Jones et al., 2006) and the Hebb effect (Sjöblom & Hughes, 2020)—are produced using order reconstruction as a response mode. As the current experiments required order reconstruction combined with the novel aspect of sequence fragment completion, pilot experiments were conducted and confirmed that the experiment design and hardware could suitably be used to test the behavioural paradigms.

Figure 6

Apparatus for MRI Experiment 1 (auditory).

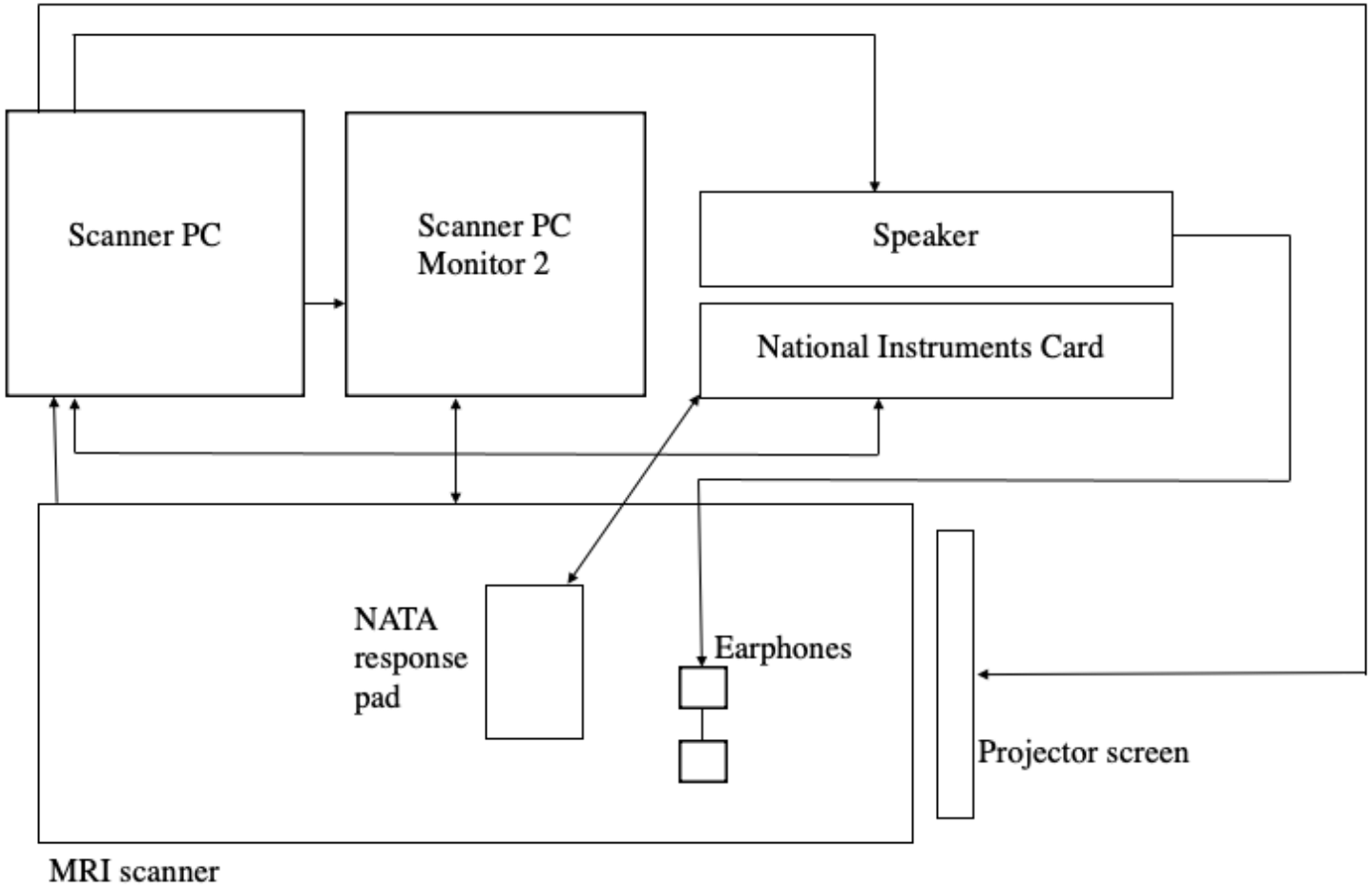
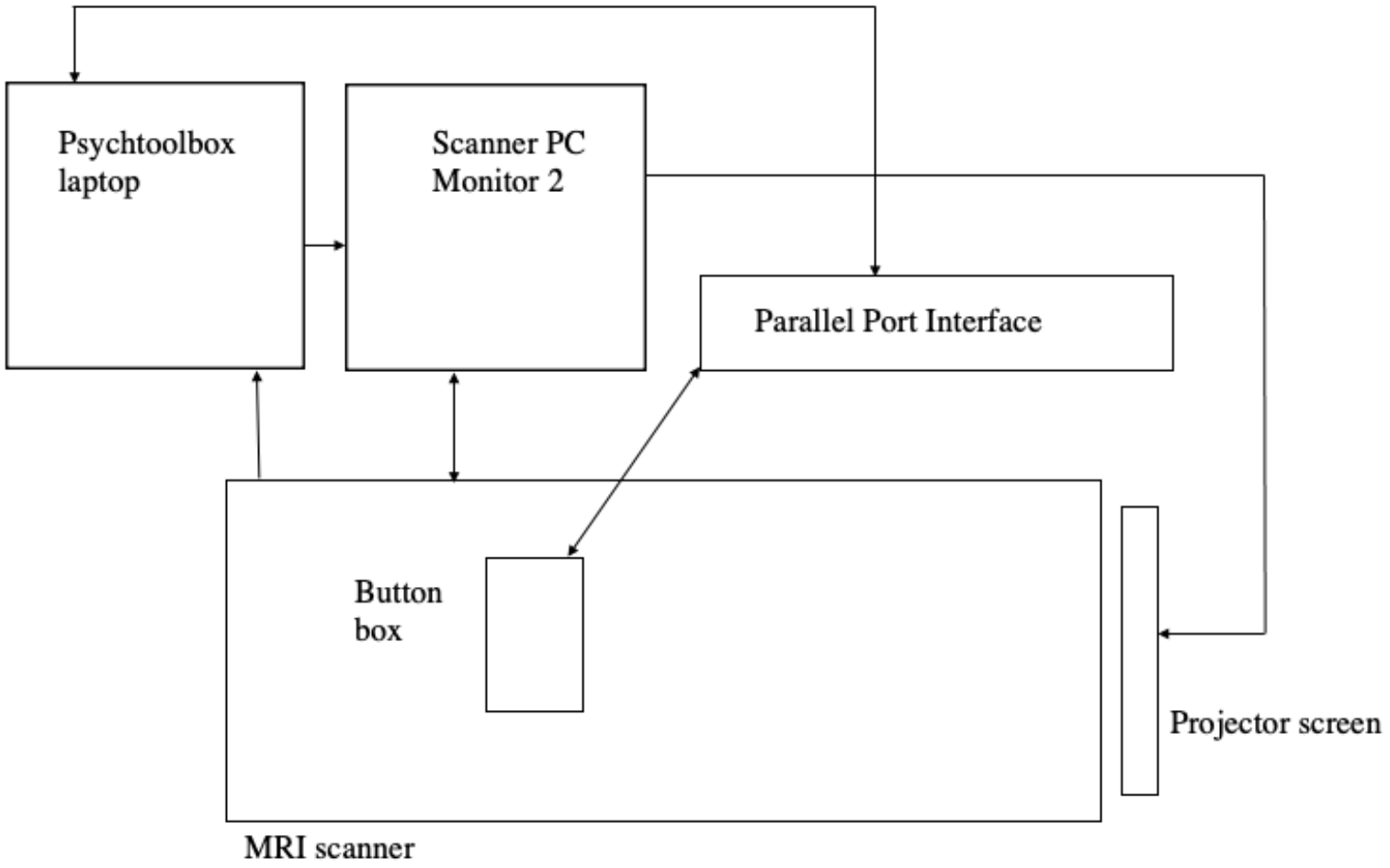


Figure 7

Apparatus for MRI Experiment 2 (visual).



2.6.2 Data Acquisition

For each subject, 800 T2* weighted Echo Planar Imaging (EPI) images were acquired using a 3 Tesla Siemens Trio Scanner (Royal Holloway, University of London). Prior to the functional scans, high resolution T1-weighted structural images were acquired at a resolution of 256 mm x 256 mm x 1 mm using an MPRAGE sequence (TR = 1900s; TE = 3.03s; flip angle = 11 degrees). The T2* EPI acquisition and experiment began following T1 equilibration. The field of view covered the whole brain: 40 oblique slices oriented (roughly) 30 degrees from transversal to coronal; field of view =192 mm x 192 mm; voxel size 3 mm x 3 mm x 3 mm; TR = 3s; TE = 30s; flip angle = 85 degrees. The TR was set to 3 s as this produces reasonable levels of signal-to-noise and can accommodate reasonable trial lengths based on the temporal jitter requirements (see Figure 4A). Following the functional scans, field maps were acquired.

2.6.3 Procedure

Subjects completed consent forms and information sheets prior to having the details of the experiment and procedure explained to them. During acquisition of the T1 structural scan, subjects completed practice trials where task performance was observed from the MRI control room to confirm their understanding of the task. The functional run began once verbal consent to continue was gained via an intercom connecting the control room to the scanner. Subjects were instructed prior to entering the scanner that communication of consent was the only overt vocalisation required as part of the experimental procedure. It was reinforced that no other vocalisations or orofacial movements (e.g., silent mouthing) should occur. Subjects were also informed that a grey screen would indicate the end of the experimental session. They

were instructed to remain still as field maps were acquired for another 2 minutes following functional acquisitions.

2.7 Functional MRI data analysis

2.7.1 Pre-processing

MRI data were pre-processed using an Automatic Analysis (aa) pipeline (Cusack et al., 2015) in MATLAB and SPM12 (Wellcome Trust Centre for Functional Neuroimaging, London, www.fil.ion.ucl.ac.uk/spm). Automatic Analysis is an open-source framework where an analysis pipeline comprises a series of modules, each of which performs a specific task.

2.7.1.1 Normalisation and registration.

Each subject's structural and functional images were normalised to Montreal Neurological Institute (MNI) space (Friston et al., 1995a,b) using Diffeomorphic Anatomical Registration Through Exponentiated Lie algebra (DARTEL) (Ashburner, 2007; Taylor et al., 2017) and is used to achieve more accurate inter-subject registration and spatial normalisation of images. In doing so, it is more likely that significant activations will be present and anatomical localisation of activations will be highly accurate. Three rigid-body translations and rotations (x, y and z planes) were estimated at the realignment stage.

2.7.1.2 Segmentation, warping and co-registration

First, the structural images of Experiment 1 subjects (n = 20) and Experiment 2 subjects (n = 21) were rigid-body registered (i.e., aligned) with the T1-weighted template in the canonical MNI space. The template was derived from icbm_avg_152_t1_tal_lin.mnc. To correct for intensity bias, the structural images were reduced to 2 mm isotropic resolution and smoothed with an 8 mm FWHM

Gaussian filter. Next, structural images were segmented and used to create a study template. Then a 12-degrees-of-freedom, affine linear transformation (i.e. affine transformation matrix) was calculated and saved, which co-registered the study template to the MNI space. Structural images were co-registered to the study template using non-linear transformation (i.e. warp field), which was also saved. The affine transformation matrix was then applied to the subjects' structural images in the study space, so that the resultant image was in MNI space.

2.7.1.3 Realignment, slice timing, extended co-registration and smoothing.

The EPI images were spatially re-aligned to correct for head motion, whilst temporal realignment corrected for different slice acquisition times. Subsequently, EPI images were then co-registered to the subjects' structural images. The non-linear warp field transformation and affine transformation matrix (used for structural images) were used to normalise the EPI images so that the resultant functional images were in MNI space. These images were then resliced to 3 x 3 x 3 mm voxels, and a Gaussian kernel of 8 mm was applied to spatially smooth the images to conform to the spatial smoothness assumption of the Gaussian Random Field Theory, as implemented in SPM12 (Friston et al., 1995). The fieldmaps acquired at the end of the functional run were used to correct for any spatial distortion. A flowchart of the processing pipeline can be found in Appendix B.

2.7.2 Quality Assurance Diagnostics of fMRI data

The quality of the fMRI data of each individual subject was assessed with a dedicated module within the aa pipeline. This module produced a report prior to motion correction of EPI images and again following motion correction. The resultant graphs from motion correction are shown in Appendix D and enabled identification of where significant changes in intensity occurred, either within slices or across volumes. The

scaled variance graph takes the 4-dimensional data set and calculates the variance across volumes. Spikes here could be due to scanner artefacts or subject movement; ideally, the graph would show minimal spikes with small amplitude (<20). In the slice variance per volume graph, all slices should have similar variance across volumes. Slices with higher variance are usually the lower ones in the brain as they are affected most by signal drop out and cardiac-induced pulsation of the brain stem. The scaled mean voxel intensity displays the mean of the BOLD-signal for each volume. Here, we would want to observe a small range in signal intensity; however, any fluctuations due to breathing and any low-frequency sinusoidal periodicity or linear drift can be removed by means of a high-pass filter during modelling. The slice variance graph shows the maximum, mean, and minimum variance across all voxels within a particular slice and shows this for each of the 40 slices acquired. The scaled mean slice intensity is a variation of scaled mean voxel intensity. It averages the signal in all voxels per slice and assesses signal intensity changes where slices are shown on the Y axis. From the latter slice intensity graph, the FFT of slice intensity (log transformed) is calculated. FFT in this instance converts the signal from the time-series domain to a representation in the frequency domain.

The Discrete Fourier transform (DFT) decomposes the sequences of values from the signal into components of different frequencies displayed in the graph. These highlight dominant frequencies of the BOLD-signal intensity and would display any possible periodicity due to either movement, physiological or MRI scanner artefacts. The displacement graph shows translation (mm) and the rotation (degrees) per volume, and the final plot shows scan-to-scan displacement by adding scan-to-scan translations and rotations in their original units. All of the aforementioned graphs were used to assess the quality of the fMRI data for each individual subject.

Based on the motion correction parameters (translation and displacement), scaled variance per volume, scan to scan displacement and FFT, outliers in the data were observed and some subjects were consequently removed from second-level group analyses. The exclusionary criteria were based on displacement and translation of the functional scans over 2mm simultaneous with large spikes in the scan-to-scan displacement graph. There was no numerical upper limit (mm and/or scaled degrees) set for exclusion of subjects based on spikes in the scan-to-scan displacement; however, this graph was assessed based on the frequency and height of spikes in comparison to the variation in displacement across the acquisition of all 800 volumes. The scaled variance per volume was similarly assessed in terms of frequency and height of spike and the frequency of BOLD-signal intensity in the log transformed FFT graph was indicative of any artefacts which may have highlighting outliers in the dataset based on displacement and translations relative to the rest of the subjects rather than to a fixed threshold. The default setting for MATLAB box plots displaying outliers is a whisker range of 1.5 x the interquartile range (IQR) and therefore anything more than 1.5 IQR beyond the 3rd quartile is considered an outlier.

2.7.3 First-level, General Linear Model

The overall modelling strategy was to enable observation of activity time-locked to the temporal delay (event 2) independently from other events (events 1 and 3). The temporal jitters (see Section 2.2.4) between each of the three events ("Stimulus presentation", "Temporal delay" and "Recall"), and between scan onset and event 1 onset, enabled isolation of the BOLD-signal time locked to each event. Assessing the degree to which activity within each voxel changes in relation to task requirements procedures can be done through a mass univariate GLM. The observed fMRI time-series in any given voxel, Y, can be approximated by the product of model factors, X,

and a scaling factor β, of which the linear combination across these approximates to Y plus a noise term ε. This can be expressed in simple matrix notation as $Y = X\beta + \varepsilon$. The statistical analysis of fMRI data results in statistics indicating evidence against the null hypothesis of no effect at each voxel. Assessing the data at each voxel is parametric and specific forms of probability distribution are assumed for the data.

For each subject, a GLM was constructed comprising 20 experimental regressors (see Appendix C). Each regressor represented multiple, repeating events of the same trial type. The temporal properties of each trial (see Section 2.2.3) were designed to allow modelling of individual trial phases. Each square wave function defined by the onset and duration of an event was convolved with a canonical hemodynamic response function to create a regressor in the GLM. The canonical HRF is the sum of two gamma functions which model the peak and undershoot of the BOLD response. Whilst the canonical HRF is often considered the most appropriate for GLMs due to its simplicity, it is possible to flexibly model parameters that capture canonical HRF variations as the shape of the BOLD response is known to vary across different regions of the brain—due to variations in vasculature—across subjects and trials (Duann et al., 2002). Partial derivatives of the canonical HRF include a temporal derivative that can capture differences in the latency of the peak response and a dispersion derivative that can capture differences in the duration of the peak response. Indeed, it has been shown that significant variability is captured by these derivatives individually and when combined suggesting that they are sufficient in capturing experimental variance (Friston et al., 1998, Henson et al., 2001). The derivatives however, were not used in the current experimental design. As explained in Sections 2.2.4 and 2.5, the low correlation coefficients between experimental regressors

indicated that little shared variance was present between experimental regressors and that using the canonical HRF was suitable in this instance.

The 20 regressors included in the first-level model were events 1, 2 and 3 of conditions “Similar-Go”, “Dissimilar-Go,” “Filler” and “Hebb” and events 1 and 2 of “Similar-No-Go” and “Dissimilar-No-Go” as there was no event 3 “Recall” for “No-Go” trials. Additionally, regressors for “Hebb” event 2 and “Filler” event 2 with parametric modulations of first order and second order polynomial expansions were also included.

Table 2
Table of regressors (shaded blue).

	Event 1 Presentation	Event 2 Temporal delay	Event 2 Temporal delay First-order polynomial expansion	Event 2 Temporal delay Second-order polynomial expansion	Event 3 Recall
Experiment A					
Similar-Go					
Similar-No-Go					
Dissimilar-Go					
Dissimilar- No-Go					
Experiment B					
Filler					
Hebb					

A first order polynomial expansion captures variance in the BOLD-signal that changes linearly over time and the second order polynomial expansion captures quadratic changes. Including two polynomial expansions results in two parametrically modulated regressors (expanded into linear and quadratic terms). In this instance, orthogonalising the Nth order term with respect to Nth-1 assigns shared variance to the lower order terms. Linear trends were therefore orthogonalised with respect to the onset and the quadratic trend is orthogonalized with respect to the linear trend and the

onset. Although behavioural learning effects are often demonstrated by upward linear trends, it is possible that the concomitant decreases of the BOLD-signal may not follow pre-empted trends. Therefore, when hypothesising a change in the BOLD-signal over time, including both first-order and second-order polynomial regressors, allows flexible modelling without having to make assumptions about the shape of the signal amplitude. Using parametric modulators in an attempt to observe a learning effect limits the model and reduces inferential uncertainty. If, therefore, a decrease in BOLD-signal across the course of Experiments 1B and 2B is observed in "Hebb" (repeating) trials whilst activity for "Filler" (non-repeating) trials remains consistent, a reliable difference between Hebb and Filler and hence learning of a repeating sequence can be deduced. The same polynomial expansions for Filler and Hebb were included into a full factorial analysis of variance (ANOVA). Although no change over time was hypothesised for Filler (non-repeating) sequences, expansions were included in the second-level analysis to detect any possible changes that may have occurred. Six head-motion regressors calculated at the realignment stage representing the three rigid-body translations and rotations (x, y and z planes) were also included in each GLM to capture residual movement-related artefacts. A first-level single subject design matrix can be found in Appendix C.

2.7.4 Second-level analyses

GLMs were estimated at the first-level of analysis and SPM{t} contrasts were run such that contrast images were obtained for each column in the design matrix. These were then incorporated into separate second-level full-factorial ANOVAs (per study: auditory and visual). The aims were firstly to observe activation of regions associated with perceptual-motor processing during presentation and rehearsal of stimuli and secondly, by using an identical trial design with a repeating sequence, assess the effect of time on rehearsal-related activity.

Table 3

Table highlighting first-level SPM {t} images (shaded blue) used in second-level contrasts. Contrast "i" was conducted using {t} images from Experiment A, contrasts "ii-iv" were conducted on {t} images from Experiments A and B whilst contrast "v" was conducted on data from {t} images from Experiment B.

	Contrast	i) Temporal delay (Go) > Temporal delay (No-Go)	ii)Presentation > Temporal delay	iii) Temporal delay > Presentation	iv) Presentation ∣ Temporal delay (conjunction)	V) Hebb (repeating) < Filler (non-repeating)
Events						
Similar-Go Event 1						
Similar-Go Event 2						
Dissimilar-Go Event 1						
Dissimilar-Go Event-2						
Filler-Go Event-1						
Filler-Go Event2						
Filler-Go Event 2 (linear expansion)						
Hebb-Go Event 2						
Hebb-Go Event 2 (linear expansion)						

The following contrasts were conducted (contrasts i-iv and were collapsed across the factor of phonological similarity):

i) *Temporal delay (Go) > Temporal delay (No-Go),* to identify activation specific to the temporal delay where subjects were expected to continue rehearsing the sequence.

ii) *Presentation > Temporal delay (Go),* to identify activation specific to the presentation event. Events from Hebb trials were not included as the BOLD-signal during 'Stimulus presentation' of the repeated Hebb sequences may also have begun to decrease across the course of the experiment and thus affected the results of this contrast.

iii) *Temporal delay (Go) > Presentation,* to identify whether any regions were active during the temporal delay in the absence of experimental stimuli, over and above activation in the presentation event.

iv) *Presentation | Temporal delay (Go)* (conjunction), to identify activation that was common to both presentation and temporal delay events.

v) *Repeating (Hebb) < Non-repeating (Filler),* to identify graded changes of the BOLD-signal during rehearsal of a repeating sequence in comparison to a non-repeating sequence. The model for Experiment B of each study included event 2 ('Temporal delay' phase) for 'Filler' and 'Hebb' and additional regressors for both trial types. These regressors were parametric

modulations of first and second order polynomial expansions (linear and quadratic respectively).

To assess the extent to which the BOLD-signal decreased as behavioural accuracy of the repeating sequence increased across the course of Experiments 1B and 2B, different ways in which the data could be modelled were considered. The first and most obvious test of whether a correlate of the behavioural Hebb sequence learning effect (main effect of List-type) was observable in the functional imaging data was a {T} contrast Hebb < Filler. To pre-empt the results, this contrast showed no reliable activations, and it became obvious that to assess the haemodynamic changes evoked by a repeating sequence, the model needed to account for the behavioural learning effect more specifically. As results of the behavioural analysis demonstrated a main effect of List-type, it therefore seemed logical to use the values from the behavioural data as parametric modulators. The group-level mean for Filler and Hebb collapsed across the seventeen cycles was used as weighting in SPM12 at the second-level stage of analysis. However, this again revealed no significant activations. Next therefore, a group level average for both trial-types per cycle (cf. behavioural graphs in empirical chapters) was included as a covariate; however, this also showed no results. Instead of attempting to account for group behaviour explicitly, therefore, the approach settled on was to use general trends by calculating the difference between the degree of the slope for the Hebb and Filler curves using this as a parametric modulation. However, a single value was not enough. Instead, SPM polynomial expansions (linear and quadratic) of Hebb and Filler were used. This allowed for flexibility in the model, necessary when investigating group level learning data, as individual differences are marked. The simplest contrast to test for Hebb sequence learning was to use regressors

Hebb (linear decrease of BOLD-signal) vs. Filler (static BOLD-signal) as behavioural accuracy for the Filler (non-repeating) condition was not expected to change over the course of the experiment but that for the Hebb (repeating) condition was.

2.8 Anatomical Methods and Data Interpretation

2.8.1 Anatomical Localisation

Peak co-ordinates were localised to gross anatomical landmarks on the canonical brain of the MRI series and verified on T1 structural images of the subjects included that had been normalised to the reference space of the MNI template when activations were ambiguous on the MNI template. The locations of peak haemodynamic activity are reported in terms of MNI co-ordinates and these are attributed to sulcal and gyral landmarks using terminology from the Duvernoy atlas (1999) general neuroanatomical reference. For the cerebellum, the atlas of Schmahmann et al., (1999) as well as the nomenclature of Larsell (1972) were used when discussing gross cerebellar anatomy. Cytoarchitectonic probability values of peak co-ordinates reported in results tables (Chapters III and IV) are based on the probabilistically defined cytoarchitectonic maps in the SPM Anatomy Toolbox (Eickhoff et al., 2005). Such results were only included if the SPM Anatomy Toolbox reported values at 50% or above. A probability lower than 50% indicated that there was a greater probability that the identified region was not an established cytoarchitectonic region.

Assigning a gross anatomical landmark and cytoarchitectonic area (if applicable) to results was completed for each experiment and co-ordinates were compared for proximal locations. For any areas of ambiguity, group activation was overlaid on each individual subject's T1 structural image and a tally was taken. Five sets of co-ordinates across all results were removed from results tables as voxels fell

into white matter. These activations were not considered further as they were considered to be noise. Given that the normalisation parameters were checked and found to be correct, it is clear that this noise in the data has manifested as activation.

2.8.2 Masks for Small Volume Correction

Cerebellar, prefrontal and premotor cortex activity for {T} contrasts during second-level group analysis was tested using small volume correction, $p < .01$ (SVC) (Worsley et al., 1996) as the experimental hypotheses were based on these specific regions. Instead of FWE-correcting for all voxels across the entire brain, using SVC's enabled analysis more specific to the hypotheses and increased the likelihood of observing specific effects in those regions. The cerebellar mask was generated using the SUIT toolbox in SPM12 (Diedrichsen, 2006; Diedrichsen et al., 2009; Diedrichsen & Zotow, 2015). The prefrontal cortex mask was manually constructed by the author using FSLeyes and the Harvard-Oxford Cortical Atlas (Desikan et al., 2006; Frazier et al., 2005; Goldstein et al., 2007; Makris et al., 2006). Changing the opacity and colour rendering of the cortical atlas allowed for the clear identification of the anatomical boundaries of areas included in the atlas. Using the location function simultaneously with the label provided, the following areas anterior to the precentral sulcus were selected to be included in the mask: frontal pole, insular cortex, superior frontal gyrus, middle frontal gyrus, inferior frontal gyrus, precentral gyrus, frontal medial cortex, juxtapositional lobule cortex (formerly supplementary motor cortex), subcallosal cortex, paracingulate gyrus, cingulate gyrus (anterior division), frontal orbital cortex and frontal operculum cortex.

CHAPTER III: AUDITORY-VERBAL SERIAL SHORT-TERM MEMORY AND LEARNING

3.1 Introduction

The understanding of verbal serial short-term memory has arguably been dominated by an account that posits a fractionation between a specialised passive phonological store and an active articulatory rehearsal process that supports that store (Salamé & Baddeley, 1982; Vallar & Baddeley, 1984). Of particular interest here is that this dominance has extended beyond cognitive psychology to cognitive neuroscientific investigations of verbal short-term memory (e.g., Awh et al., 1996; Chen & Desmond, 2005a,b; Desmond et al., 1997; Henson et al., 2000; Paulesu et al., 1993; Peterburs et al., 2016, 2019; Salmon et al., 1996; Smith et al., 1996). The general consensus from these studies is that the active articulatory rehearsal process can be mapped onto motor-related areas of the brain such as the premotor cortex, cerebellar lobule VI and Crus I (Chen & Desmond, 2005a,b; Desmond et al., 1997) whilst the passive phonological store can be localised to the inferior parietal lobe, though where within this lobe has remained elusive (Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996).

In recent years, however, cognitive-behavioural research has suggested that performance in verbal serial short-term memory tasks can be explained in terms of the formation of an articulatory plan for task-relevant motor output and processes involved in perceptual organisation, without having to invoke a distinct passive phonological store (the *perceptual-motor account*; Hughes & Marsh, 2017; Hughes et al., 2009, 2016; Jones et al., 2006, 2004; Macken et al., 2016; Sjöblom & Hughes, 2020). In light

of this, a reappraisal of the cognitive neuroscience of verbal short-term memory seems warranted: The present fMRI study examines whether the neural underpinnings of verbal serial short-term memory performance, as well as the long-term learning of a verbal sequence—the supposed evolved function of the phonological store (Baddeley et al., 1998)—can be understood in terms of the action of brain regions or networks that are involved in articulatory planning and those involved in perceptual or perceptual-motor mapping processes.

Recent doubts concerning the existence of a specialised phonological store have revolved around what has long been cited as its chief behavioural hallmark: the phonological similarity effect, the poorer serial recall of a list of phonologically similar items (e.g., 'B', 'C', 'D'…) compared to a list of phonologically dissimilar items (e.g., 'F', 'K', 'R'…). Importantly, this phonological similarity effect is found even when articulatory rehearsal is impeded by articulatory suppression so long as the list is presented auditorily rather than visually. This has led to the view that there must, therefore, be a passive phonological store separate from articulatory processes (Baddeley et al., 1984). Auditory-verbal items gain obligatory access to the store, hence the survival of the phonological similarity effect under suppression with such items, whereas visual-verbal items require the articulatory rehearsal process to gain access to the store, consequently resulting in the disappearance of the similarity effect under suppression with visual presentation. Under closer examination, however, the residual phonological similarity effect under suppression with auditory sequences occurs only at recency (the end of the list), which has been acknowledged as falling outside the extent to which a phonological store construct could account for the effect (Baddeley, 1986) and is, instead, readily explained as the product of passive acoustic-based perceptual organisation (Jones et al., 2004; Nicholls & Jones, 2002). Thus,

articulatory rehearsal is a prerequisite for the phonological similarity effect regardless of presentation-modality, hence locating the effect within the articulatory rehearsal process, not a passive phonological store. The striking resemblance between the kinds of errors made in the serial recall of phonologically similar lists and naturally occurring speech-planning errors (e.g., Acheson & MacDonald, 2009; Ellis, 1980) provides further independent evidence for an articulatory-planning basis to the phonological similarity effect. Such evidence has been instrumental in the development of an alternative perceptual-motor approach that denies the existence of a specialised store and which sees verbal serial short-memory performance, as well as long-term verbal sequence learning, as being parasitic on motor planning and perceptual processes (Hughes et al., 2009, 2016; Jones et al., 2006, 2004; Macken et al., 2016; Sjöblom & Hughes, 2020, Chapter 1, Section 1.4).

The cognitive-behavioural work that has served to challenge the phonological store construct—and that has led to the alternative perceptual-motor approach—motivates the present re-evaluation of the cognitive neuroscience of verbal short-term memory. Previous cognitive neuroscientific studies of verbal short-term memory have for the most part focused on localising the putative phonological store, with some localising it along the supramarginal gyrus (Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996), others locating it in the intraparietal sulcus (IPS; Smith et al., 1996) and still others localising it to the posterior parietal cortex (Awh et al., 1996; Jonides et al., 1997). Moreover, activation of regions across the parietal cortex varied *within* an investigation comprising several verbal tasks that were all meant to reveal the location of the phonological store (Henson et al., 2000). Thus, one immediate difficulty with the endeavour to localise the phonological store has been that no consensus has yet been reached as to its location, other than the rather general claim

that the parietal cortex is involved. A further difficulty is that the interpretations of the neuroimaging data are often at odds with the fairly specific predictions that follow from the phonological loop model (cf. Chapter 1, Section 1.5): For instance, the proposal that the store is located in the posterior parietal cortex (Awh et al., 1996; Jonides et al., 1997) can be questioned on the grounds that auditory stimuli are meant to gain obligatory access to the store and yet the posterior parietal cortex is not active during passive listening to auditory-verbal stimuli (Binder et al., 2000). Instead, auditory processing is typically associated with bilateral activation in superior temporal cortices, not the parietal lobe. A rather different problem arises in relation to the localisation of the store in the IPS insofar as this region is found to be active in a visual-spatial and not just a verbal short-term memory task (Smith et al., 1996). As such, it is difficult to accept the claim that the IPS is the site of a store specialised for verbal representations. In addition, the localisations of the store in the supramarginal gyrus (Paulesu et al., 1993; Salmon et al., 1996) are more superior to the Sylvian-parieto-temporal region typically associated with short-term memory deficits resulting from focal brain lesions (Shallice & Warrington, 1980; Shallice & Vallar, 1990; Vallar & Papagno, 1995;).

It is notable also that in early attempts to localise the phonological store, the particular task-phase during which parietal activation was observed was not specified (Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996). As noted by others (Buchsbaum & D'Esposito, 2019; Morey et al., 2019), such information could be highly diagnostic theoretically because the phonological model predicts quite clearly that any neural signature of the specialised phonological store should be evident during all phases of a verbal short-term memory task: the presentation of the sequence as items access the store and begin to be rehearsed in

articulatory form (to offset the decay of the items therein), during any retention interval when articulatory-based refreshing of items would be expected to continue, and finally at the response stage when items are being retrieved from the store. Results of later studies that did use event-related experimental designs in order to examine activations relating to different phases of a short-term memory task are also, therefore, inconsistent with the predictions of the phonological loop model: Chein and Fiez (2001) observed activation in the left inferior parietal cortex only during stimulus presentation while Chen and Desmond (2005b) only observed activation in this region during a retention interval. A further study found that only the right inferior parietal cortex was active during presentation (Marvel & Desmond, 2010a). More specific localisation revealed that bilateral activation in the intraparietal sulcus was observed only during a retention interval (Sakai et al., 2002) and activation in the left supramarginal gyrus only during retrieval (Marvel & Desmond, 2010a).

What has been highly consistent across studies, however, is the interpretation that activation of the premotor and supplementary motor areas (BA 6) along with the inferior frontal gyrus (BA 44), and lobules HVI and HVIIA (Crus I) of the cerebellum indicate the operation of an articulatory processes in verbal short-term memory (Awh et al., 1996; Chen & Desmond, 2005a,b; Henson et al., 2000; Marvel & Desmond, 2010a, Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996). While it should be noted that not all these regions were found to be active in all of these studies, the weight of evidence (cf. Chapter 1, Table 1) nonetheless indicates much better agreement as to which regions are involved in articulatory rehearsal as compared with the case in relation to the putatively separate function of phonological storage. Consistent with the view that the cerebellum may support rehearsal processes is evidence of anatomical connections in non-human primates between the primary and

premotor cortices with lobule HVI (Kelly & Strick, 2003, Lu et al., 2007). However, with the studies being set within the phonological loop framework, these motor regions of the cerebellum were seen as merely supporting the phonological store located in the parietal cortex. It was suggested that a store was supported by cerebellar lobules VIIB/VIII (Chen & Desmond, 2005a,b; Desmond et al., 1997; Kirschen et al., 2005; Kirschen et al., 2010). Such an interpretation is, however, contrary to neuroanatomical evidence demonstrating that lobules HVIIB and HVIII project directly to the primary and premotor cortices (Kelly & Strick, 2003, Lu et al., 2007). Instead, the approach in this thesis focuses on the mechanics of an articulatory rehearsal process via the transmission of information between prefrontal and premotor regions of the cortex with cerebellar lobules to which they are connected. Evidence of non-human primate anatomical connections, resting-state functional connectivity in humans and cerebellar cytoarchitecture (cf. Chapter 1, Section 1.6) have implicated the cerebellum as a key structure in the acquisition of motor skills and suggest a reinterpretation of what has previously been considered a peripheral process to verbal short-term memory proper as the very means by which a verbal sequence is maintained and reproduced.

The primary aim of the current study, therefore, was to examine whether (auditory-) verbal serial-short term memory performance can be accounted for in terms of the activation of brain systems involved in motor planning, perceptual processing or perceptual-motor mapping without having to invoke a distinct phonological-store proposed to be located in the inferior parietal lobe. In the first part of the present experiment (Experiment 1A), participants were presented with a seven-letter sequence followed by a short (5 s) temporal delay, following which they were required to complete a re-presented fragment of the sequence (four of the seven letters). The letters were either phonologically similar or dissimilar to one another. Replicating the classic

phonological similarity effect would provide independent corroboration that subjects had engaged in articulatory planning (from the standpoint of the perceptual-motor account) or utilised the phonological store (from the perspective of the phonological loop model). The phonological similarity manipulation was used simultaneously with contrasting trials where subjects were instructed as to whether recall—and thus retention of to-be-remembered items—or rest was required.

Comparison of haemodynamic activation during the temporal delay of trials termed 'Go' and 'No-Go' trials (regardless of phonological similarity) were hypothesised to reveal activation in the prefrontal cortex (area 46) on the basis that the region has consistently been implicated in supporting action-planning via the selection and maintenance of goals and rules at higher-levels (Badre & D'Esposito, 2009; Lashley, 1951; Miller et al., 1960). Simultaneous activation of cerebellar lobule HVIIA (Crus I and Crus II) – to which area 46 has anatomical connections – was also hypothesised (Kelly & Strick, 2003; Lu et al., 2007). As the prefrontal cortex directly projects to the premotor cortex (Orioli & Strick, 1989; Lu et al., 1994), activation of the premotor cortex (BA 6) and its cerebellar targets—lobules HIV-HVI (Kelly & Strick, 2003, Lu et al., 2007) were also predicted. This activation may suggest that neural activity in the cerebellum related to motor commands and feedforward error-predictions during covert articulation. It has been suggested that Crus I and Crus II contain forward models that predict the responses of premotor cortex to higher-level commands issued by the prefrontal cortex (Ramnani, 2006). These predictions are proposed to modify actions thus enabling automatic planning and outputs. As Lu et al. (2007) did not specify whether tracers were injected in the primary or premotor cortex, activation of lobules HVIIB and HVIII may also reflect motor-related activity in line with projections observed by Kelly and Strick (2003). Finally, activation of the inferior

frontal gyrus (BA 44, pars opercularis) is also predicted as the region is proposed to be involved in the construction of articulatory representations and the mapping of intended speech to orofacial musculature (Papoutsi et al., 2009; Hickok & Poeppel, 2004; Price, 2012; Rauschecker & Scott, 2009). Taken together, activations of these regions across the delay period of 'Go' trials would be interpreted as being indicative of an articulatory rehearsal process.

A second key set of predictions concerned which areas would be recruited during presentation compared to the temporal delay phase. The perceptual-motor account predicts that an appeal to articulatory motor planning and perceptual processing should obviate the need to posit a store separate from motor and perceptual regions. Activation during sequence presentation was compared to that during the temporal delay and was predicted to reveal regions involved in perceptual, or perceptual-motor mapping processes that would not be observed consistently across both task phases. These regions included the auditory cortex and planum temporale. Activation of any other regions are expected to be explainable by recourse to perceptual or motoric processing. A key consideration in the perceptual-motor account then, is how perceptual processes are mapped onto a motor plan and whether evidence of auditory-motor mapping is evident in results from the present study, thus proposing an alternative view to specialised storage for verbal information. A conjunction analyses between sequence presentation and the temporal delay was also conducted to assess any similarities in activations across these task phases. As it is assumed that articulatory rehearsal begins during the presentation of to-be-remembered items and continues across a temporal delay, activation of all regions implicated in rehearsal should consistently be active across both task phases whilst no region indicative of a phonological store should be observed. Alternatively, if a phonological store does

indeed exist, then the neural correlate should be observed in results of a conjunction analysis as items enter and are refreshed within the store (cf. Chein & Fiez, 2001) and a contrast comparing activation across a temporal delay where sequence retention is required, to that where it is not. When taken together, activations in results across all contrasts are expected to be explicable by recourse to motor planning and perceptual processing, rather than the operation of a phonological store that is distinct from motor, speech and language processes.

Another key issue studied here was the neural basis of the long-term learning of a verbal sequence. Such learning has been described as the evolved function of the phonological store (Baddeley et al., 1998) but here it was examined whether, instead, such learning may again be understood more parsimoniously in terms of (alterations in) the operation of brain regions associated with motor planning (cf. Sjöblom & Hughes, 2020). Experiment 1B, therefore, employed the Hebb sequence learning paradigm in which one sequence in a block of serial recall trials is intermittently repeated in amongst non-repeating 'filler' sequences (Hebb, 1961). Traditionally, increased behavioural accuracy for a repeating sequence presented every three to four trials is observed across the course of an experiment— and was used to assess whether the signature of articulatory rehearsal during the temporal delay observed in Experiment A, changed over time as a function of learning.

As behavioural performance increases, decreases in neural excitability across the cerebellum are predicted to arise from error signals correcting for discrepancies between intended and actual movements (Albus, 1971). To further support this, research has demonstrated that high frequency neural activity at the start of learning declines to background levels as learning progresses (Gilbert & Thach, 1977). Cerebellar networks demonstrating evidence of long-term synaptic plasticity indicate

that experience-dependent adaptive learning processes are a salient feature of cerebellar function (Ekerot & Kano, 1985; Ito et al., 2014; Robinson, 1976; Thach, 1998). Control theory suggests that experience-dependent learning—the decline of error-related activity during motor learning—results in more accurate predictions from forward models stored in the cerebellum (Kawato & Wolpert, 1998; Ramnani, 2006; Wolpert & Kawato, 1998; Wolpert & Miall, 1996). This results in a transition of control from cortical to cerebellar areas as actions become increasingly automatic (Ramnani, 2014) and a breadth of research has shown decreases in cerebellar activity as motor sequence learning occurs (Doyon et al., 2002, 2003; Grafton et al., 1994; Imamizu et al., 2000; Penhune & Doyon, 2002; Tzvi et al., 2014). Consistent with evidence of cortico-cerebellar connections and theories of automaticity and cerebellar learning, an excitability decrease (Albus, 1971) was predicted in cerebellar lobules HIV-HVI, HVIIB and HVIII that are connected to the primary and premotor cortices (Kelly & Strick, 2003; Lu et al., 2007) throughout Experiment B, in line with the notion that forward models in cerebellar lobules support the increasingly automatic production of a repeating verbal sequence (i.e. motor skill learning).

3.2 Method

3.2.1 Subjects

Having obtained ethical approval from Royal Holloway's Ethics Committee (Code: 418-2017-28-17-05-PDJT002), twenty-four subjects were recruited from Royal Holloway, University of London's Experiment Management System in exchange for a small honorarium and took part in both sub-experiments (Experiments 1A and 1B). Following the experiment, each subject also received a structural image of their brain as an additional token of appreciation. Exclusion criteria were based on factors from the quality assurance report, details of which can be found in Chapter 2, Section 2.7.2.

Four subjects were excluded from the fMRI analyses due to excessive movement during scan acquisitions. The fMRI results presented in this chapter are therefore based on the data from twenty subjects (Mean age = 23 years, SD = 3.6 years, female = 13, male = 7). The behavioural results to be reported below, however, were based on the data from all 24 subjects. All subjects were right-handed and reported being psychologically and neurologically healthy.

3.2.2 Design

Whilst details of the design and procedure were provided in Chapter 2, the main features are highlighted again here. Experiment 1A comprised four trial-types, with 12 trials of each: Phonologically similar—Go, Phonologically similar—No-go, Phonologically dissimilar—Go and Phonologically dissimilar—No-go. For each subject, these four trial-types were mixed in a 48-trial block in a pseudo-randomised fashion. The requirement for sequence recall was indicated following sequence presentation with a 'Go' signal while a 'No-Go' signal informed subjects that they would not need to recall the sequence (see Chapter 2, Section 2.2.3, Figures 4A and 4B). A 'Go' trial comprised three main task phases: stimulus presentation, a 5 s temporal delay and recall. The recall phase was not included in 'No rehearsal' trials. In both instances, subjects were auditorily presented with a sequence of letters and then visually presented with a 'Go' or 'No-Go' signal which indicated whether recall (and thus retention) would be required. In the event of a 'Go' trial, and following a temporal delay, subjects were presented with a recall screen where they were required to complete a fragment of the sequence in the correct serial order using items that had been removed from their original serial position and presented in a mixed order underneath the fragment. If subjects reached the maximum allotted time for recall, a

new trial began. Following the temporal delay in 'No-go' trials, the presentation of sequence items in a new trial began as 'No-go' trials did not require recall.

Experiment 1B succeeded Experiment 1A with no break in between. It consisted of 60 trials, all of which required recall (i.e., were all 'Go' trials), and thus it was assumed that subjects would be engaging in rehearsal during the temporal delay phase of each and every trial. There were forty-three non-repeating 'Filler' sequences and 17 instances of the same (i.e., repeating) 'Hebb' sequence that all involved pseudo-random orderings of the seven phonologically similar letters used in Experiment 1A.

3.2.3 Procedure

Subjects completed consent and safety forms following an explanation of the experimental procedure. Each subject received a new pair of earbuds for health and safety reasons, which were fitted onto the MRI-safe earphones. Prior to completing the practice trials, subjects were shown each of the individual letters on the screen and simultaneously heard the spoken letter over the earphones. This was to reinforce the accuracy with which participants would be able to identify each to-be-remembered item, especially given the experimental environment. During the T1 anatomical scan, subjects completed four practice trials, where the volume was set to a comfortable level for each subject based on their feedback and adjusted accordingly. Following the completion of the practice trials, subjects were asked via the MRI intercom system whether the volume was comfortable and suitable for continuing with the experiment proper or if any adjustments needed to be made. The volume was increased only once for one subject as a result of this procedure.

3.3 Results

3.3.1 Behavioural Results

3.3.1.1 Experiment 1A (Independent test of engagement in articulatory rehearsal/phonological storage).

The percentage of items correctly recalled was significantly lower for lists with phonologically similar items (M = 54.2% SE = 2.83) than for lists with phonologically dissimilar items (M = 66.4% SE = 3.38), $t(23) = -4.486$, $p < .001$. This replication of the classic phonological similarity effect verifies from a perceptual-motor perspective that subjects had engaged in articulatory serial rehearsal and, from the perspective of the phonological store theory, would be taken as verification that the phonological store was being utilised. More generally also, the replication of this canonical effect in verbal serial short-term memory indicates that the novel adaptation of the verbal serial recall task used here was suitable for studying the neural basis of verbal serial short-term memory and learning.

3.3.1.2 Experiment 1B (Hebb Sequence Learning).

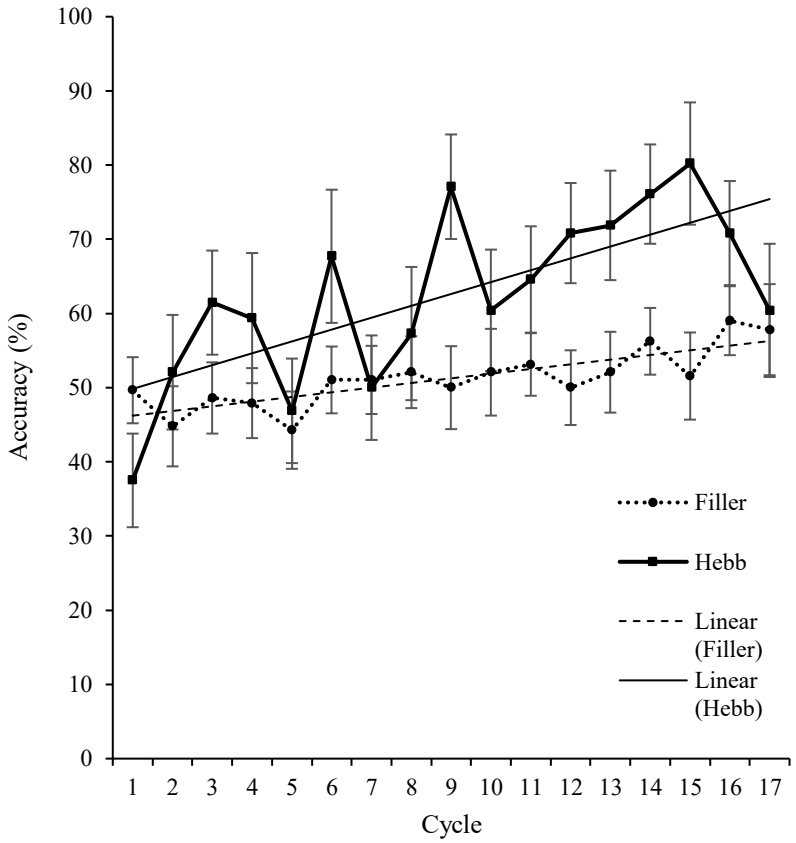
Figure 8 shows the percentage of items recalled correctly when presented with a Filler sequence and when presented with the repeating Hebb sequence as a function of cycle. It is clearly evident that, overall (collapsing over cycles), recall was markedly better on the Hebb trials (M = 62.8% SE = 3.4) than on the Filler trials (M = 51.3% SE = 2.6), thereby replicating the classic Hebb effect (Hebb, 1961). Recall of the Hebb sequence also increased generally across the 17 cycles (ranging from M =37.5% in cycle 1 to M = 80.2% at cycle 15)[4] whereas there was relatively little improvement across cycles for Filler trials. A 2 (List-type) × 17 (Cycle) repeated measures analysis of variance

[4] It is unclear why recall of the Hebb sequence suddenly dropped during the last two cycles but as no such effect was found under similar conditions in Experiment 2B and given that the data in the Hebb condition were generally rather noisy, no functional significance will be attached to this.

(ANOVA) confirmed the reliability of the pattern of results just described: There was a main effect of List-type (Hebb, Filler), $F(1, 23) = 20.28$, $MSE = 1323.4$, $p < .001$, $\eta_p^2 = .469$, a main effect of Cycle, $F(16, 368) = 3.126$, $MSE = 737.144$, $p < .001$, $\eta_p^2 = .120$, and a reliable List-type by Cycle interaction, $F(16, 368) = 1.893$, $MSE = 701.7$, $p = .02$, $\eta_p^2 = .076$.

Figure 8

Percentage accuracy for Filler and Hebb sequences as a function of cycle, including linear trendlines, in Experiment 1B.



3.3.2 fMRI results

Throughout the remainder of the Results section, several references are made to parameter estimate graphs accompanying images of activation on axial, coronal or sagittal slices of the MNI152 template structural T1 scan. The graphs demonstrate the direction and amplitude of activation in a voxel (see co-ordinates in each figure legend) for regressors included in a given statistical test. All regressors included in second-level analyses are listed in Chapter 2, Section 2.7.4, Table 3. Tables 4-7 show comprehensive lists of results from each of the contrasts conducted but the reporting in the text will focus on those relevant to the experimental hypotheses or those deemed potentially important for interpretation even if they were not predicted. It should be noted that any localisation of co-ordinates to BA 10 in the middle frontal gyrus that were not localised as FP1 or FP2 (Bludau et al., 2015) using the SPM Anatomy Toolbox, were done so based on the localisation of area 9/46 (Petrides & Pandya, 1999, 2002) as BA 10 sits anteriorly to area 9/46. Small volume corrections (SVC; Chapter 2, Section 2.8.2) were applied to results of the *Temporal delay (Go) > Temporal delay (No-Go)* and *Repeating (Hebb) > Non-repeating (Filler)* contrasts. Results observed under SVC or both FWE-correction and SVC are denoted by symbols throughout Tables 4 and 7; see table headings for detail. The column header "BA" in all results tables refers to Brodmann areas indicated either by the SPM Anatomy toolbox (only included if cytoarchitectonic probability > 50%) or using the sulcal and gyral anatomy of Duvernoy (1999; cf. Chapter 2, Section 2.8.1 for rationale). Data are presented separately in parameter estimate graphs for scans from trials with phonologically similar and dissimilar lists but there was no expectation of differences as a function of phonological similarity and hence contrasts were collapsed across this factor.

3.3.2.1 Experiment 1: Temporal delay (Go) > Temporal delay (No-Go).

This contrast was designed to reveal regions active during the temporal delay period of 'Go' trials—when subjects were expected to keep rehearsing the sequence in preparation for recall—as compared to the temporal delay period of 'No-Go' trials (when subjects were to merely await the start of the next trial). In line with the hypotheses, a region in the prefrontal cortex along the middle frontal gyrus (BA 46, Figure 9B) and premotor cortex in the pre-central gyrus (BA 6, Figure 9A) were active during the articulatory rehearsal of the sequence (see Table 4). More specifically cytoarchitectonic subdivisions within the pre-SMA: 6MR (Amunts et al., 2019; Amunts & Zilles, 2015) and 6D3 (Sigl et al., 2019) were revealed. Cerebellar lobules HIV, HVI, HVIIB were also active and the pattern of results shown in the parameter estimates (Figures 10A-10C) is similar to that in the premotor cortex (BA 6, Figure 9A) and prefrontal cortex in the middle frontal gyrus (BA 46, Figure 9B).

Figure 9

Parameter estimates and images from the results of the Temporal delay (Go) > Temporal delay (No-Go) {T} contrast in Experiment 1. A: Sagittal slice of MNI152 template showing *activation in premotor cortex and parameter estimate in area 6mr (-6, -3, 57). B: Axial slice of MNI152 template showing activation in middle frontal gyrus, area 46 (-33, 48, 18).*

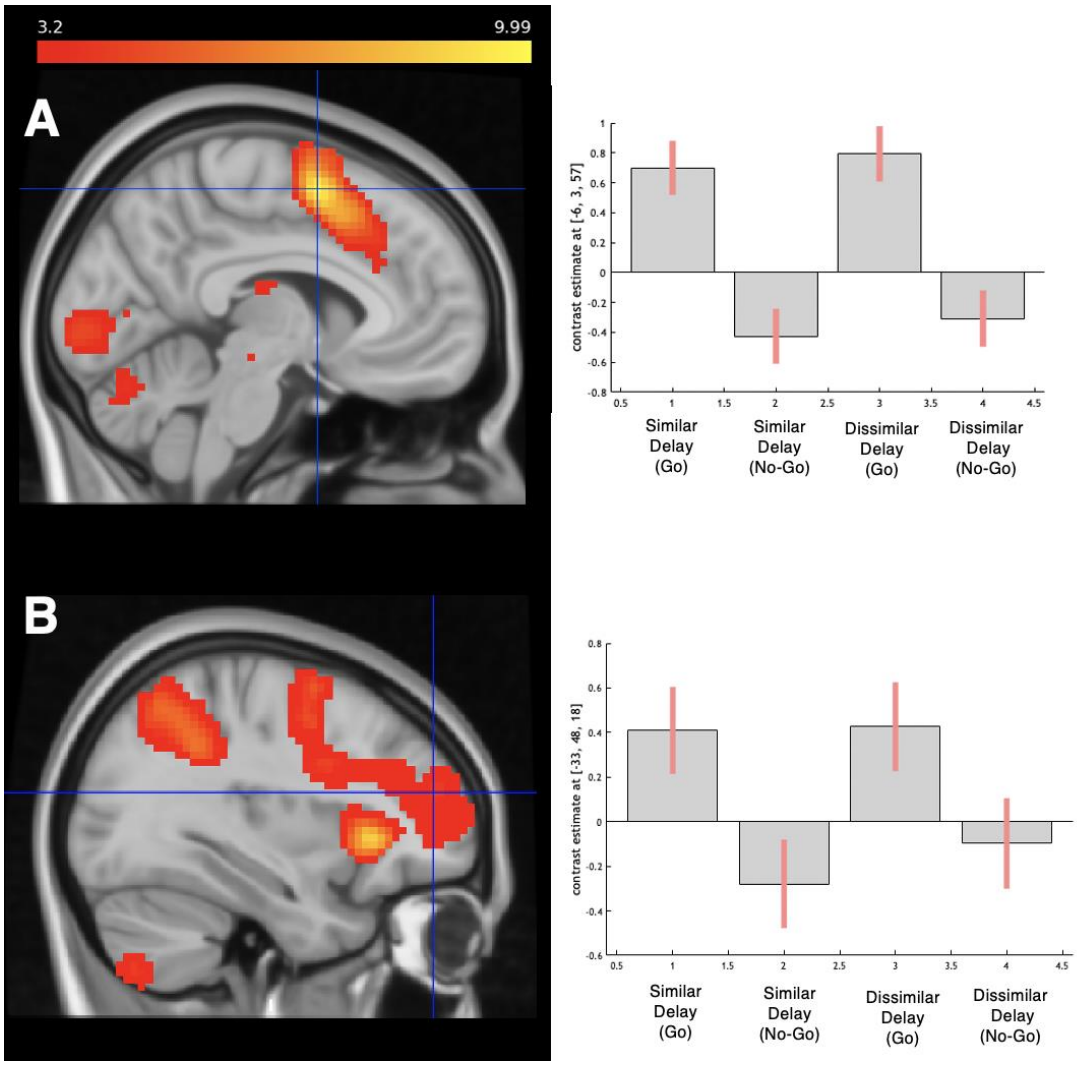


Figure 10

Parameter estimates and images from results of Temporal delay (Go) > Temporal delay (No-Go) {T} contrast in Experiment 1. Coronal slices of MNI152 template showing activation in the cerebellar lobule A: IV, B: HVI, C: HVIIB, D: HVIIA (Crus I) and E: HVIIA (Crus II).

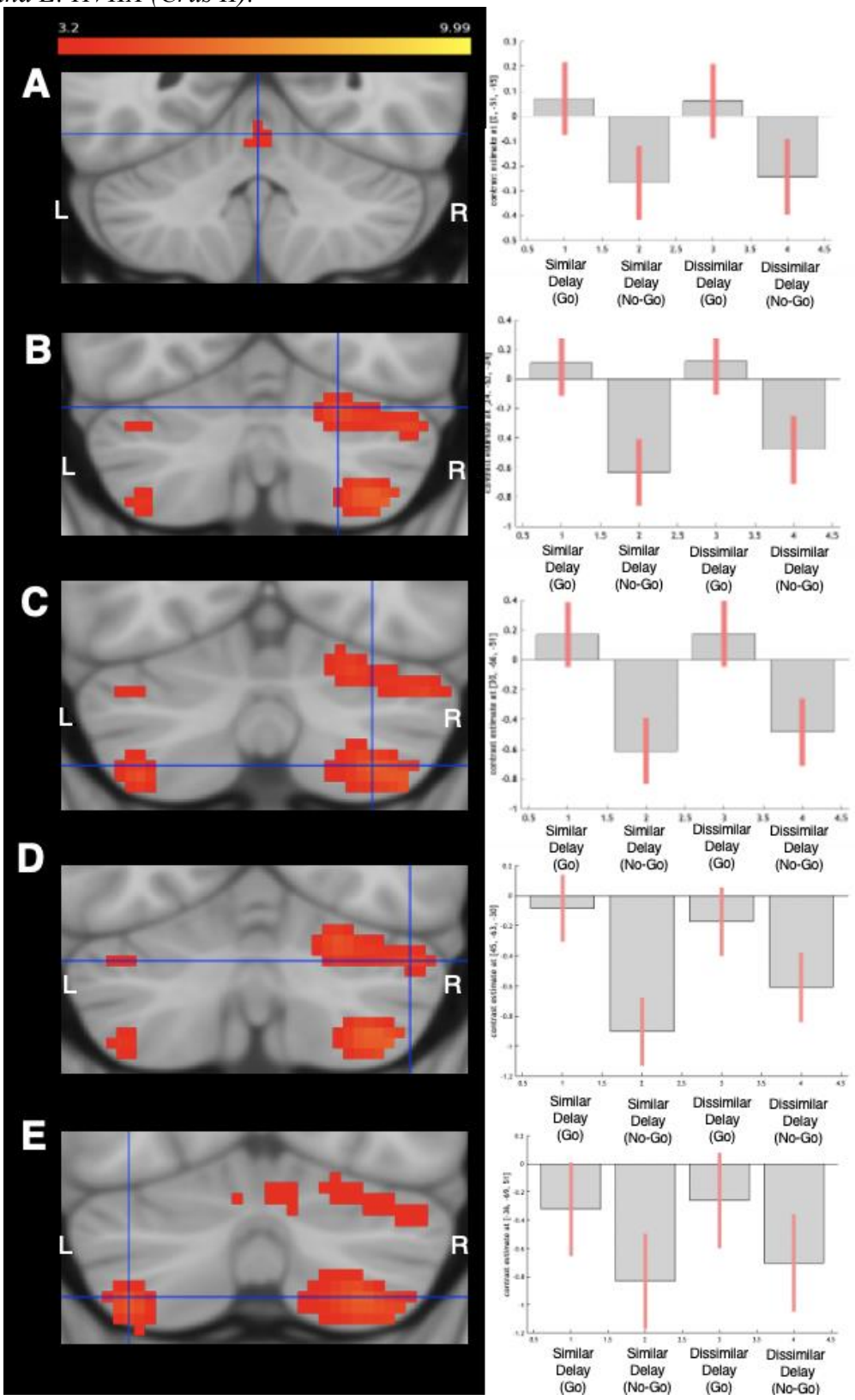


Table 4

*Results of a paired t-test, Temporal delay (Go) > Temporal delay (No-Go) {T} contrast (FWE-corrected for multiple comparisons, p < .05; *Prefrontal and Premotor cortices SVC; **Whole Cerebellum SVC; † Prefrontal, & Premotor cortices SVC and FWE-corrected for multiple comparisons, p < .05; ‡ Whole Cerebellum SVC and FWE-corrected for multiple comparisons, p < .05).*

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Frontal Lobe							
Prefrontal cortex							
Middle frontal gyrus†	6.99	6.13	45	39	27	9	
Middle frontal gyrus*	6.16	5.53	-42	21	30	9 & 46	
Middle frontal gyrus*	5.97	5.39	-48	27	33	9	
Frontopolar gyrus†	5.33	4.9	27	57	-6	10	
Frontopolar gyrus*	5.1	4.71	27	51	-3	10	
Middle frontal gyrus*	5.04	4.67	-36	12	30	9	
Middle frontal gyrus†	5.01	4.64	-33	48	18	46	
Middle frontal gyrus*	4.99	4.63	-36	33	24	46	
Frontopolar gyrus	4.91	4.56	33	48	0	10	
Orbital gyrus†	4.88	4.54	-21	51	-12	11	
Middle frontal gyrus*	4.44	4.18	33	45	12	46	
Frontopolar gyrus*	4.33	4.08	-33	57	6	10	
Middle frontal gyrus*	4.05	3.84	48	15	33	9	
Premotor cortex							
Superior frontal gyrus†	9.92	Inf	-6	3	57	6MR	65.3%
Pre-central gyrus	6.57	5.83	-45	0	48	6	
Pre-central gyrus*	6.49	5.77	-48	-3	54	6	
Pre-central gyrus*	6.43	5.73	-54	0	39	6	
Pre-central sulcus*	5.74	5.21	-27	-3	51	6D3	51.4%
Pre-central sulcus*	5.62	5.12	-30	0	60	6	
Pre-central gyrus*	5.54	5.06	-36	6	30	6	
Pre-central sulcus*	5.53	5.06	-42	0	27	6	
Pre-central sulcus*	4.61	4.32	30	0	54	6	
Pre-central gyrus*	4.26	4.02	30	9	60	6	
Pre-central sulcus*	4.14	3.92	30	-6	48	6	
Inferior pre-central	4.01	3.81	39	6	30	6	
Temporal Lobe							
Insular cortex							
Insula†	8.79	7.28	-30	24	0	Id7	72.9%
Insula†	7.81	6.67	30	24	-3		
Insula*	6.57	5.83	-45	0	48	Id7	72.9%
Cingulate cortex							
Paracingulate gyrus†	7.68	6.59	0	15	48	32	
Cingulate sulcus*	4.79	4.46	12	18	33	24	
Anterior Cingulate	4.31	4.07	-9	27	27	24	
Paracingulate gyrus	4.79	4.46	12	18	33	33	

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Parietal Lobe							
Intraparietal sulcus	6.9	6.06	33	-63	51	hIP3	50.1%
Intraparietal sulcus	6.82	6	30	-60	39	hIP6	51.4%
Intraparietal sulcus	6.57	5.83	-39	-48	42	7	
Intraparietal sulcus	6.2	5.56	-30	-63	51	7	
Intraparietal sulcus	6.12	5.5	-27	-57	42	7	
Occipital Lobe							
Calcarine sulcus	4.74	4.42	-12	-90	-3	hOc1, V1	57.9%
Cerebellum							
Lobule HVIIB**	5.33	4.9	30	-66	-51		
Lobule HVIIB**	5.33	4.9	27	-72	-51		
Lobule HVIIB	5.33	4.9	30	-66	-51		
Crus II**	5.08	4.7	-36	-69	-51		
Lobule HVI‡	4.88	4.54	24	-63	-24		
Lobule HVI**	4.64	4.34	6	-75	-24		
Crus I**	4.54	4.26	45	-63	-30		
Vermis Lobule IV**	3.56	3.41	0	-51	-15		
Crus I**	3.5	3.36	-39	-66	-30		
Vermis Lobule IV**	3.3	3.18	3	-54	-21		
Lobule HVI**	3.29	3.17	-27	-60	-30		

3.3.2.2 Experiment 1: Presentation > Temporal delay (Go).

The aim of this contrast was to assess the prediction that areas specifically involved in perceptual or perceptual-motor mapping processes would be active over and above the motor planning processes that were expected to be engaged throughout both presentation and the temporal delay phases. In line with these predictions, activation was observed in the regions of the premotor cortex (BA 6) that were not active during results of the conjunction analysis (Section 3.3.2.3, Table 6), across the superior temporal lobe (BA 22) and in the primary auditory cortex (BA 41). Activation in the posterior superior temporal gyrus (MNI: -57, -21, 0) is likely to be that of the planum temporale. When overlaying group activation on the MNI152 template (see Figure 11), it was unclear as to where the co-ordinates (MNI: -57, -24, 12) were located in terms of gross anatomy. Given this, group activation was overlaid on each individual subject's anatomical scan and it was deduced that the co-ordinates were located within

the transverse temporal sulcus (as reported in Table 5) for eighteen out of twenty subjects.

Figure 11

Parameter estimates and images from the results of Presentation > Temporal delay (Go) {T} contrast in Experiment 1. Sagittal slice of MNI152 template showing activation in the auditory cortex and parameter estimate in the posterior superior temporal cortex, (MNI: -48, -39, 3).

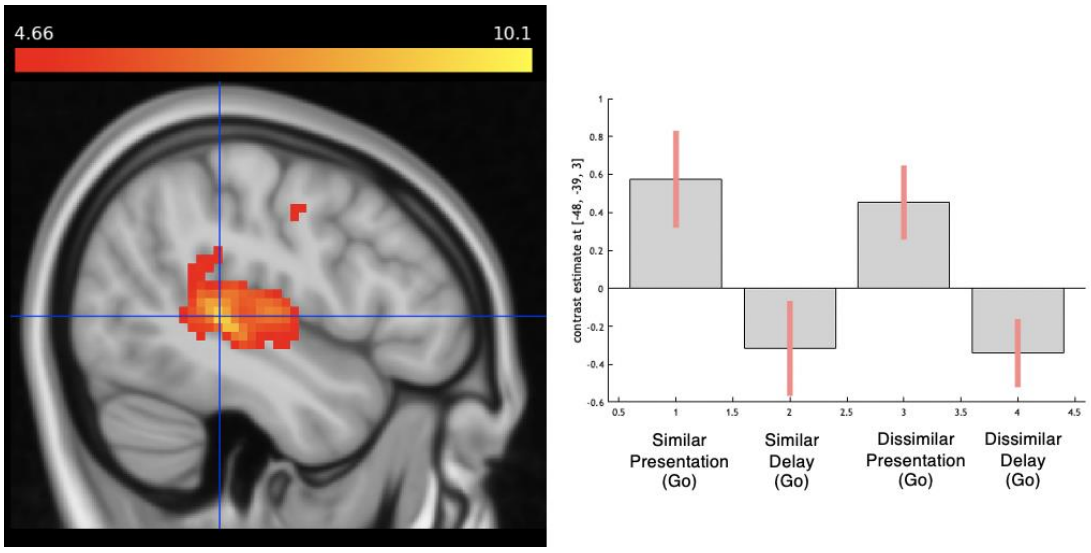


Table 5

Results of the factorial analysis of variance (ANOVA) of the Presentation > Temporal delay (Go) {T} contrast, FWE-corrected for multiple comparisons, p < .05.

Gross Anatomy	T	Z	Co-ordinates (x. y, z)			BA	Probabilistic Cytoarchitecture (if available)
Frontal Lobe							
Superior frontal gyrus	5.92	5.52	0	0	66	6	
Premotor cortex							
Pre-central gyrus	6.32	5.84	-51	-9	42	6	
Pre-central gyrus	6.06	5.64	-57	0	24	6	
Pre-central gyrus	4.69	4.48	51	-6	42	6	
Temporal Lobe							
Transverse temporal	8.95	7.78	63	-33	15	41	
Transverse temporal	8.79	7.68	-57	-24	12	41	
Superior temporal gyrus	8.43	7.43	-57	-21	0	22	
Superior temporal gyrus	8.11	7.2	66	-30	3	22	

3.3.2.3 Experiment 1: Presentation | Temporal delay (Go) (Conjunction).

A conjunction analysis was performed to assess whether predicted activation of the indicative of an articulatory rehearsal process—was consistent across both task phases or whether a region consistent with previous localisations of a phonological store could be observed. Critically, no parietal regions were observed in the results of the conjunction (see Table 6). Instead, the results showed that areas of the superior and middle frontal gyri (BA 9) and premotor cortex (BA 6)—specifically regions of the pre-SMA—were significantly active. The parameter estimate graph in Figure 12 shows activity in BA 6 suggesting that the key process common to both task phases is articulatory rehearsal.

Figure 12

Parameter estimates and images from the results of the Presentation | Temporal delay (Go) {T} conjunction in Experiment 1. Sagittal slice of MNI152 template showing activation in premotor cortex and parameter estimate in (MNI: -48, -3, 54).

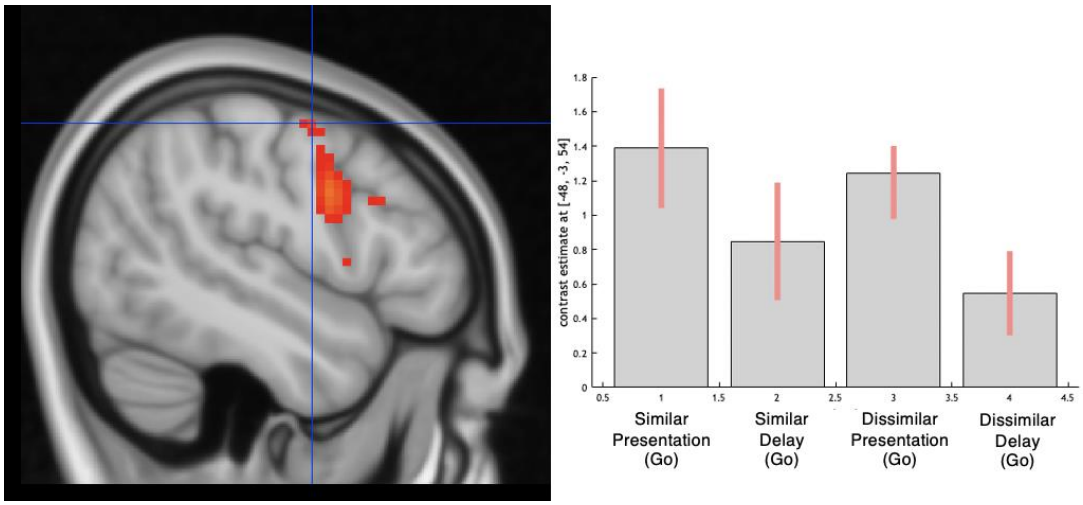


Table 6

Results of a factorial analysis of variance (ANOVA) of the Presentation | Temporal delay (Go) {T} conjunction, FWE-corrected for multiple comparisons, p < .05.

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Frontal Lobe							
Prefrontal Cortex							
Middle frontal gyrus	5.82	5.44	-42	21	27	9	
Superior frontal gyrus	4.83	4.6	-54	24	33	9	
Superior frontal gyrus	4.82	4.59	39	45	36	9	
Superior frontal gyrus	4.78	4.56	42	39	30	9	
Premotor Cortex							
Pre-central gyrus	7.26	6.58	-9	12	45	6	
Pre-central gyrus	6.91	6.31	-6	0	57	6	
Pre-central gyrus	6.29	5.82	-48	3	30	6	
Pre-central gyrus	5.97	5.56	-57	0	42	6	
Superior frontal sulcus	5.57	5.23	-21	-3	54	6d3	53.6%
Pre-central gyrus	5.47	5.15	-48	-3	54	6	
Pre-central gyrus	5.14	4.87	30	-6	51	6	
Pre-central gyrus	4.81	4.58	48	3	27	6	
Basal Ganglia							
Pallidum	7.24	6.56	-12	3	-3		
Mediodorsal thalamic nucleus	5.23	4.94	-9	-15	6		
Insular Cortex							
Insula	8.73	7.64	-30	21	3	Id7	65.8%
Insula	8	7.12	30	21	0		
Short Insular gyrus	5.88	5.49	-45	9	6		
Cingulate Cortex							
Paracingulate gyrus	6.88	6.29	9	15	45	32	
Occipital Lobe							
Lingual gyrus	4.84	4.61	-15	-90	-3	18	
Cerebellum							
Vermis Lobule IX	4.8	4.58	0	-54	-30		83.2%

3.3.2.4 Experiment 1: Repeating (Hebb) < Non-repeating (Filler).

This contrast was conducted to test the hypothesis that long-term learning of a repeating sequence, as demonstrated by the significant behavioural main effect of list-type reported in Section 3.3.1.2, would be reflected in a cycle-by-cycle decrease in the BOLD-signal as compared to the signal for non-repeating sequences. The results of the *Repeating* < *Non-repeating* contrast showed that cerebellar lobules HVI and HVIIB, observed to be active in Experiment 1A (*Temporal delay [Go]* > *Temporal [No-Go]*), were significantly active in the activation map of this contrast. On closer inspection, however, the parameter estimate graphs (Figures 13A-13F) show that learning was not sequence-specific as predicted but instead task-general. Parameter estimate graphs of the activation in lobules HV and HVIIIA (Figures 13A and 13F respectively) demonstrate similar patterns of activity to those of lobules HVI and HVIIB (Figures 13B and 13E respectively) but the former lobules were not found to be active in results of any contrast in Experiment 1A either FWE-corrected, SVC with a whole cerebellar mask, or at a lower threshold ($p < .01$). Furthermore, Figure 13 shows two coronal slices with almost identical activation patterns for repeating and non-repeating sequences across various regions of the cerebellar cortex. Thus, although Hebb sequence learning was observed in the behavioural results, no sequence learning-related changes were detected in the fMRI data.

Figure 13

Coronal slices of MNI152 template showing activation across the cerebellum in lobules HIV, HV, HVI, Crus I, Crus II and HVIIIB in both the Filler (light blue) and Hebb (dark blue) conditions, in Experiment 1. Parameter estimate plots showing activity evoked in Hebb Linear < Filler {T} contrast. A: Lobule HV, B: Lobule HVI, C: HVIIA (Crus I) , D: HVIIA (Crus II), E: Lobule HVIIB, F: Lobule HVIIIA.

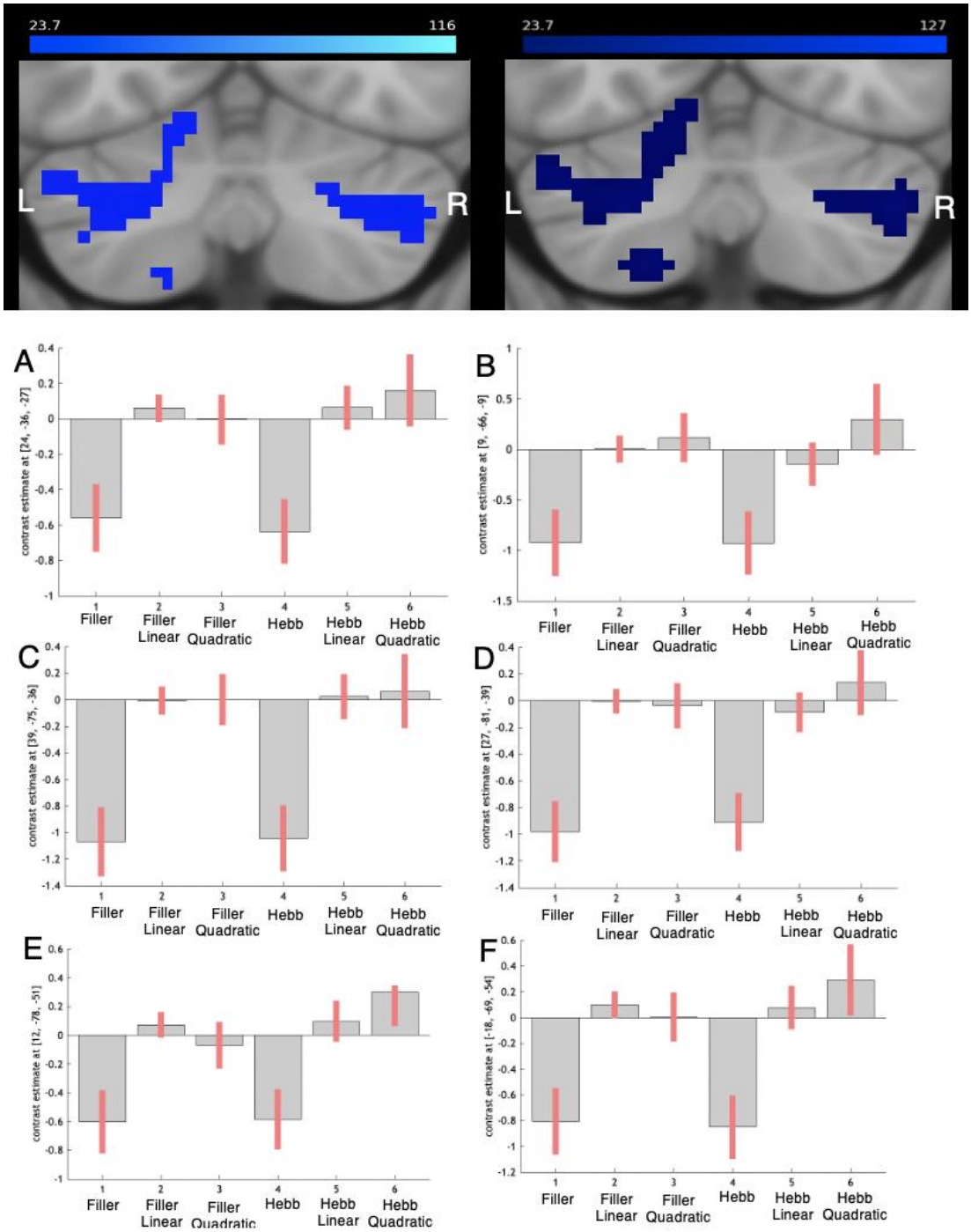


Table 7

*Results of a factorial analysis of variance (ANOVA), Repeating (Hebb) < Non-repeating (Filler) {T} contrast (FWE-corrected for multiple comparisons, p<.05; *Prefrontal and Premotor cortices SVC; **Whole Cerebellum SVC; † Prefrontal, & Premotor cortices SVC and FWE-corrected for multiple comparisons, p< .05; ‡ Whole Cerebellum SVC and FWE-corrected for multiple comparisons, p <.05)*

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Frontal Lobe							
Prefrontal cortex							
Posterior Orbital gyrus†	6.1	5.6	24	18	-18	Fo3	54.60%
Orbital sulcus†	6.02	5.6	-33	33	-18	11	
Frontomarginal gyrus*	5.79	5.4	-36	36	-15	11	
Orbital sulcus†	5.64	5.2	33	36	-18	11	
Orbital gyrus†	5.07	4.8	39	21	-18	11	
Orbital gyrus†	4.92	4.6	42	45	-15	11	
Orbital gyrus†	4.77	4.5	45	42	-18	11	
Pre-central gyrus*	4.71	4.4	-57	-3	9	44	
Superior frontal gyrus*	4.5	4.3	-36	12	45	8	
Pre-central gyrus*	4.49	4.3	-63	0	12	44	
Superior frontal gyrus*	4.45	4.2	-15	30	48	9	
Inferior frontal gyrus*	4.44	4.2	54	21	-3	45	
Middle frontal gyrus*	4.44	4.2	42	9	45	8	
Frontomarginal gyrus*	4.2	4.0	-3	48	-24	11	
Frontal Medial cortex*	4.13	3.9	-9	54	-9	Fp2	66.00%
Inferior frontal gyrus*	4.1	3.9	48	33	-3	45	
Inferior frontal gyrus*	4	3.8	54	27	6	45	65.60%
Inferior frontal gyrus*	3.92	3.7	57	12	0	44	
Superior frontal gyrus*	3.88	3.7	12	39	48	9	
Frontal pole*	3.87	3.7	-6	57	-18	Fp1	59.10%
Superior frontal gyrus*	3.76	3.6	9	36	51	9	
Inferior frontal gyrus*	3.73	3.6	-54	24	3	45	
Frontorbital gyrus*	3.69	3.5	-51	42	-12	10	
Superior frontal gyrus*	3.65	3.5	6	39	48	9	
Frontomarginal gyrus*	3.64	3.5	27	57	-12	11	
Superior frontal gyrus*	3.59	3.4	-12	69	9	Fp1	91.70%
Superior frontal gyrus*	3.52	3.4	18	33	48	8	
Superior frontal gyrus*	3.51	3.4	6	63	24	10	
Superior frontal gyrus*	3.51	3.4	12	45	39	9	
Superior frontal gyrus*	3.46	3.3	-3	60	24	10	
Frontomarginal gyrus*	3.41	3.3	12	57	-18	11	
Medial orbitofrontal	3.41	3.3	12	18	-12	Fo2	51.40%
Superior frontal gyrus*	3.37	3.2	-9	45	39	9	
Superior frontal gyrus*	3.31	3.2	-18	42	39	9	
Superior frontal gyrus*	3.27	3.1	21	63	18	10	
Middle frontal gyrus*	3.25	3.1	30	54	21	10	

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Superior frontal gyrus*	3.25	3.1	-21	57	21	10	
Superior frontal gyrus*	3.22	3.1	-21	45	36	10	
Superior frontal gyrus*	3.21	3.1	-24	42	33	10	
Middle frontal gyrus*	3.19	3.1	33	60	-9	10	
Primary motor cortex							
Pre-central gyrus	3.52	3.4	-3	-30	78	4	
Pre-central gyrus*	3.55	3.4	6	-30	78	4a	72.50%
Premotor cortex							
Pre-central gyrus*	4.66	4.4	-54	-6	21	6	
Pre-central gyrus*	4.57	4.3	60	3	6	6	
Pre-central gyrus*	4.49	4.3	63	-3	12	6	
Pre-central gyrus*	4.49	4.3	36	-18	42	6	
Pre-central sulcus*	4.38	4.2	-48	-9	27	6	
Pre-central gyrus*	4.31	4.1	-42	-12	33	6	
Pre-central gyrus*	4.27	4.1	-12	21	57	6	
Superior frontal gyrus*	4.16	4	24	-18	63	6	
Pre-central gyrus*	3.36	3.2	-9	-12	75	6d1	50.80%
Superior frontal gyrus*	3.3	3.2	-27	-27	72	6	
Pre-central sulcus*	3.26	3.1	12	18	60	6	
Temporal Lobe							
Superior temporal	8.37	7.3	-42	-33	3	22	
Middle temporal gyrus	8.37	7.3	48	-27	9	22	
Superior temporal gyrus	8.01	7.1	60	-33	15	22	
Superior temporal gyrus	7.85	7	-57	-27	12	22	
Fusiform gyrus	7.66	6.8	60	-21	12	22	
Superior temporal gyrus	7.39	6.6	-45	-21	6	TE 1.0	54.30%
Fusiform gyrus	6.27	5.8	33	-27	15	22	
Transverse temporal	4.73	4.5	-51	-57	-21	37	
Superior temporal gyrus	4.72	4.5	-63	-18	-21	21	
Insular cortex							
Insular cortex*	6.77	6.1	-33	-24	9	Ig1	69.80%
Insular cortex*	6.42	5.9	-42	-15	3		
Circular insular sulcus*	6.13	5.6	39	-21	3		
Circular insular sulcus*	6.01	5.5	39	-21	12		
Circular insular sulcus*	5.89	5.4	42	-15	6		
Circular insular sulcus*	5.66	5.3	45	-9	3		
Insular cortex*	4.87	4.6	54	-54	-18		
Insular cortex*	3.58	3.4	33	6	9		
Circular insular sulcus*	3.41	3.3	36	3	6		
Cingulate Cortex							
Cingulate sulcus†	5.92	5.5	21	-48	33	23	57.10%
Cingulate gyrus†	5.69	5.3	-3	-12	39	24	53.60%
Cingulate gyrus	5.43	5.1	27	-51	27	23	
Cingulate gyrus*	5.31	5.0	3	-15	42	24	
Anterior cingulate gyrus	5.01	4.7	24	0	30	24	

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Cingulate sulcus	4.71	4.4	6	-39	30	23	
Anterior cingulate	4.28	4.1	0	24	18	24	
Ventral cingulate	3.84	3.7	6	18	-9	24	
Anterior cingulate	3.82	3.7	-6	48	27	32	
Anterior cingulate	3.63	3.5	-12	51	0	32	
Ventral cingulate	3.62	3.5	3	15	-6	25	82%
Anterior cingulate	3.45	3.3	-9	42	3	24	
Ventral Cingulate	3.43	3.3	-6	39	9	24	
Cingulate gyrus	3.29	3.2	-6	15	-9	25	52.30%
Parietal Lobe							
Postcentral gyrus	5.58	5.2	-3	-54	69	5	
Postcentral gyrus	5.56	5.2	21	-42	72	1,2,3	
Superior parietal cortex	5.56	5.2	24	-36	60	2	63.70%
Postcentral gyrus	5.35	5.0	-21	-42	72	1,2,3	
Superior parietal lobule	5.27	4.9	15	-48	75	5L	70.60%
Parieto-occipito fissure	4.9	4.6	-18	-33	57	1,2,3	
Parieto-occipito fissure	4.85	4.6	0	-84	39	7	
Postcentral gyrus*	4.83	4.6	21	-33	57	1,2,3	
Parieto-occipito fissure	4.81	4.5	-9	-81	36	7	
Parieto-occipito fissure	4.79	4.5	6	-81	45	7	
Superior parietal lobe*	4.76	4.5	-9	-36	45	5	
Superior parietal lobe*	4.71	4.4	12	-33	45	5Ci	78.30%
Parieto-occipito fissure	4.71	4.4	12	-81	39	7	
Postcentral gyrus*	4.65	4.4	-18	-33	60	1,2,3	
Postcentral gyrus*	4.56	4.3	-15	-33	39	5Ci	78.30%
Superior parietal	4	3.8	51	-12	48	1,2,3	
Cerebellum							
Crus II‡	6.25	5.7	-33	-75	-42		
Crus I‡	5.84	5.4	42	-66	-42		
Crus I‡	5.79	5.4	-36	-75	-33		
Crus I‡	5.77	5.3	39	-75	-36		
Crus I‡	5.72	5.3	21	-78	-33		
Crus I‡	5.71	5.3	-18	-75	-33		
Crus II**	5.38	5.0	27	-81	-39		
Lobule HVIIIB‡	5.06	4.7	21	-48	-54		
Crus I‡	4.95	4.7	-51	-60	-33		
Crus I**	4.91	4.6	-48	-57	-36		
Lobule HVIIIA‡	4.71	4.4	-18	-69	-54		
Crus I**	4.51	4.3	-27	-63	-36		
Lobule HV**	4.5	4.3	24	-36	-27		
Lobule HVIIB**	4.38	4.2	12	-78	-51		
Lobule HVIIB**	4.29	4.1	-33	-60	-54		
Lobule HVIIIB**	4.22	4.0	-21	-48	-51		
Crus I**	4.13	3.9	-39	-51	-39		
Lobule HVI**	3.25	3.1	9	-66	-9		

3.4 Discussion

The results of the current experiment involving the serial recall of sequences of auditory-verbal letter-names were in line with the perceptual-motor view that an appeal to articulatory planning and perceptual processing is sufficient to account for verbal serial short-term recall performance (Jones et al., 2004). No single region or network of regions consistent with the characteristics of a phonological store (e.g., Baddeley, 2007) was found to be active. The results fail to speak to the neural basis of verbal sequence learning, however, as no sequence-specific learning effects were observed in the fMRI data. Important for establishing the suitability of the current experiment for examining the neural basis of verbal serial short-term memory (as well as learning) was the replication, in Experiment 1A, of the phonological similarity effect. There are now strong converging lines of evidence that the phonological similarity effect reflects sub-lexical speech planning errors (Acheson & Macdonald, 2009; Ellis, 1980; MacKay, 1970; Nooteboom, 1967) and hence that it can be taken as independent evidence that subjects engaged in the articulatory rehearsal of the sequence. At the same time, from the point of view of the phonological store theory, the replication of the phonological similarity effect means that it could not readily be argued that the phonological store was not utilised in the current task for some reason (e.g., Baddeley & Larsen, 2007). That is, the experiment was capable in principle of revealing the neural basis of the operation of that store.

A key contrast of interest aimed to compare activation during the temporal delay of ‘Go’ trials to that during the temporal delay of ‘No-Go’ trials (Section 3.3.2.1). Results showed that regions in the prefrontal and premotor cortices, along cerebellar lobules connected to those regions were significantly active. Given proposals that area 46 maintains higher-order representations related to motor

commands (Ramnani, 2006), it can be suggested that activation of this region in the context of verbal sequence rehearsal reflects the maintenance of high-level representations related to motor-planning processes (Badre & D'Esposito, 2009; Lashley, 1951; Miller et al., 1960). MNI co-ordinates -33, 48, 18 and -36, 33, 24 reported in Table 4 lie within BA 46. It is highly likely that they also lie in area 9/46 given the proximity of the co-ordinates to the region outlined by Petrides and Pandya (1999, 2002). Area 9/46 was termed so on the basis that a portion of Brodmann's area 9 that lies in the middle frontal gyrus was shown to exhibit similar cytoarchitectonic properties to area 46 (Petrides & Pandya, 1999; 2002). Probabilistic representations of area 10 in the frontal pole (FP1; Bludau et al., 2014) are anterior to area 9/46 (Petrides & Pandya, 1999; 2002), thus localisations of area 9/46 were identified in relation to area 10, BA 9 and BA 46. MNI co-ordinates -42, 21, 30 lie on the border of BA 9 and BA 46 and were listed as "9 & 46" in Table 4. Other regions along the superior frontal gyrus (BA 9) were also active: The activation of BA 9 simultaneous with area 46 is not uncommon, as both areas are thought to encode abstract information and monitor information maintained over delays (Funahashi 2001; Fuster 1997; Petrides 1994). In the present study, abstract information is suggested to take the form of higher-order goal representations related to action planning and selection. However, BA 46 was not shown to be active during both presentation and the temporal delay (*Presentation | Temporal delay [Go]*). Instead, activation along the superior and middle frontal gyri in BA 9 was observed across both phases and may reflect the maintenance of such high-level representations as areas 9, 46 and 9/46 exhibit similar cytoarchitectonic properties (Petrides & Pandya, 1999).

The activation of the lateral prefrontal cortex, the pre-SMA and premotor cortex in results of the *Temporal delay (Go) > Temporal (No-Go)* (Table 4),

Presentation > Temporal delay (Go) (Table 5), *and Presentation | Temporal delay (Go)* (Table 6) contrasts are consistent with the notion of translating higher-order information generated in the prefrontal cortex—regarding action selection and planning— into motor commands for the planning of movement designed to achieve goals (Hoshi, 2008). As shown in Table 4, subdivisions of the premotor cortex (6MR, 6D3) were significantly active during rehearsal of a sequence (*Temporal delay [Go] > Temporal delay [No-Go])*. Based on cytoarchitectonic properties of the region, subdivision 6MR falls within the pre-SMA (Amunts et al., 2019; Amunts & Zilles, 2015), anterior to SMA proper.

The pre-SMA maintains reciprocal connections to areas 9 and 46 in the prefrontal cortex (Bates & Goldman-Rakic, 1993) as well as the anterior premotor cortex (Johansen-berg et al., 2004; Lu et al., 1994; Luppino et al., 1993). During tasks that involve articulatory planning, activation in the pre-SMA has been associated with higher-level planning processes (Picard & Strick, 2001; Nachev et al., 2008) necessary for learning new sequential procedures, regardless of item identity within those sequences (Hikosaka et al., 1996). It could be then, that the pre-SMA is involved in updating representations in the prefrontal cortex related to the timing of sequences (Coull et al., 2016) following the generation of motor commands in the premotor cortex. Activation of the pre-SMA has also been associated with the inhibition of vocal (as well as manual) outputs (Xue et al., 2008). Although the perceptual-motor approach does not assume that overt rehearsal is substantively different from covert rehearsal in terms of underlying processes, there are additional processes involved in overt vocalisation (e.g., movement of orofacial musculature, exhalations) that may be inhibited during covert rehearsal. This is particularly likely in the current instance as subjects were explicitly instructed not to produce any overt orofacial movements as

they often result in head motion and have the potential to disrupt the acquisition of functional imaging data.

The parameter estimate graphs relating to BA 6 (6D3 and 6MR in the pre-SMA), BA 46, and cerebellar lobules IV, HVI and HVIIB (Figures 9A, 9B, 10A, 10B and 10C) show similar activation patterns and their co-activation is consistent with the operation of a cortico-cerebellar loop (Coffmann et al., 2011; Kelly & Strick, 2003; Kreinen & Buckner, 2009; Lu et al., 2007; O'Reilly et al., 2010). In support of this, studies have shown activation of lobule HIV during silent articulation (Kawashima et al., 2000)—which has been suggested to be a result of subliminal orofacial and laryngeal muscle activity (Ackermann et al., 1998)—and the activation of lobule HVI during overt vocal responses (Hayter et al., 2007). The co-ordinates of premotor (BA 6) activation observed in the *Presentation > Temporal delay (Go)* contrast (cf. Table 5) were not found in the results of the other contrasts, suggesting that specific areas of the premotor cortex could be involved in the mapping of perceptual input to motor commands. The areas of the premotor cortex that were co-active with regions across the temporal cortex (indicative of auditory-perceptual processing) may facilitate encoding of temporal and sequential properties, specific to auditory-verbal information (Kotz & Schwartze, 2010). Activation of the prefrontal cortex and IFG were not observed in results of the *Presentation > Temporal delay (Go)* contrast, however areas of the prefrontal cortex were shown to be consistently active across both presentation and delay phases. This demonstrates that no regions in the prefrontal cortex were significantly active above and beyond those observed in the conjunction and lends further support to the idea that higher-order representations in the context of this experiment were related to motor-planning processes, constant across both task phases. It is surprising that activation of the IFG was not observed in results of the

Presentation > Temporal delay (Go) contrast, even at a lower threshold, however it may be possible that the formation of articulatory codes was dealt with by regions in the dorsal premotor cortex and the planum temporale (discussed later on). Cerebellar lobules HVI, HVIIA (Crus I and Crus II) HVIIB and HVIII were active at a lower threshold ($p < .01$)

The present interpretation of cerebellar contributions to vocal motor planning in verbal serial short-term memory differs from that offered by other researchers who have investigated cerebellar activity during verbal short-term memory performance (reviewed in Chapter 1, Section 1.6, e.g., Chen & Desmond 2005a,b; Desmond et al., 1997). For instance, Desmond et al. (1997) suggested that representations involved in articulatory rehearsal (a frontal-cerebellar loop involving lobules HVI and HVIIA) and separate representations held within a phonological store (fed by a temporo-parieto-cerebellar loop implicating cerebellar lobule HVIIB/HVIIIA) were compared via error-corrective feedback and feedforward commands from the cerebellum. This view, however, is incompatible with the original phonological-store theory. The articulatory rehearsal process was proposed to refresh items *within* the store, separate representations in the store and maintained by the rehearsal process were not suggested to be compared in some manner. This incompatibility also prompts the question as to why separate representations and mechanisms were deemed necessary by Desmond et al. (1997) and later by Chen and Desmond (2005a,b).

In recent years, studies in non-human primates have shown direct connections between cerebellar lobule HVIIB and the motor cortex (Kelly & Strick, 2003, Lu et al., 2007), whilst studies of resting-state functional connectivity in humans have suggested that this lobule is connected to the prefrontal cortex (Kreinen & Buckner, 2009; O'Reilly et al., 2010). In light of this evidence, and the perceptual-motor

interpretation of processes involved in verbal serial short-term memory, it can be suggested that the connection of lobule HVIIB to areas in the frontal lobe supports vocal-motor planning via error-feedback regarding timing and sequencing that have been shown to be imperative for speech production (Ackermann, 2008; Leggio & Molinari, 2015; Molinari et al., 2008). These feedforward commands from lobule HVIIB could therefore be relayed to the prefrontal cortex in response to higher-order representations necessary for action selection and planning, or, to the premotor cortex in response to the generation of specific preparatory motor commands. Moreover, the pattern of lobule HVIIB activation as a function of task-phase in the present experiment is particularly problematic for the view that it supports a phonological store: Activation indicative of a phonological store should be evident across all task-phases but the activation of lobule HVIIB was only observed during the temporal delay and not during presentation. These results are further supported by those of Durisko and Fiez (2010) in results from their delayed serial recall task. Although the authors report that activation of lobule HVIIB was attributed to a presentation phase, detailed examination of their statistical methods revealed that sampling of the BOLD-signal at the end of the presentation phase (2 s) and start of the rehearsal phase (4 s) were compared to baseline. This would suggest that activation of lobule HVIIB is attributed to the rehearsal of items rather than their presentation. Compounding the difficulty with ascribing a role to lobule HVIIB in the operation of a phonological store, no region that has been suggested previously to be the location of such store (supramarginal gyrus; Paulesu et al., 1993, Salmon et al., 1996) was co-active with lobule HVIIB during the temporal delay in the present experiment.

Consistencies in the activation of lobules HVI and HVIIB when task requirements necessitate vocal-motor planning further support the assumption that

these lobules support general-purpose motor planning. In the present study, the activation of lobule HVI (MNI: 6, -78, -22, *Temporal delay [Go] > Temporal [No-Go])* was in close proximity to that observed by Hayter et al. (2007) (MNI: 6, -75, -24). In that study, activation of lobules HVI and vermal VIIB were significantly active during the "Add" condition of a paced serial addition task in which subjects were required to add a presented digit to the immediately preceding digit. In a "Repeat" control condition subjects were only required to repeat each number after hearing it. Lobule HVI was active during a conjunction between Add and Repeat indicating its role in the speech motor control that would have been required in both conditions. Furthermore, Durisko and Fiez (2010) showed that activation of lobule HVI was observed during all three phases in their 'memory condition' (presentation, maintenance and recall) as well as during covert speech, covert tapping, overt speech and overt tapping. In particular, similar areas of lobule HVI found to be active during the temporal delay in the present experiment (MNI: 24, -63, -24) were observed by Durisko and Fiez (2010) during covert speech (repeating "the"; MNI: 25, -58, -26) and the 'recall condition' where subjects were required to overtly recall a sequence vocally (MNI: 25, -64, -26). Additionally, activation of lobule HVI observed during the presentation of individual letters in their verbal memory condition (MNI: -25, -54, -29) and during overt finger tapping (MNI: -32, -58, -26) were also in close proximity to a cluster of activations observed across the temporal delay in the present experiment (MNI: -27, -60, -30). Thus, taken together, these results suggest that areas within lobule HVI contribute to both covert and overt sequencing of actions in repetitive and serial order that are not specific to particular effectors.

It was hypothesised that Crus I and Crus II of cerebellar lobule HVIIA would also be active during the rehearsal of sequences across a temporal delay given

suggestions that they may contain forward models that predict the responses of premotor cortex to commands issued by the prefrontal cortex (Ramnani, 2006). However, the parameter estimate graphs (Figures 10D and 10E) indicated that an excitability decrease of lobule HVIIA occurred during the temporal delay. This could suggest that forward models in Crus I and Crus II were acquired early on, perhaps towards the end of the presentation phase, prior to the temporal delay. No cerebellar activation was observed during presentation (*Presentation* > *Temporal delay* [*Go*]) at FWE-corrected level. At a lower threshold *($p < .01$)*, however, cerebellar lobules HVI, HVIIB and HVIII as well as bilateral Crus I and Crus II were found to be active. These results suggest that the operation of prefrontal and premotor projecting lobules was indeed required during the presentation phase, but that higher order representations related to motor planning may have been learned towards the end of the presentation phase prior to the temporal delay. The parameter estimate graphs of Crus I and Crus II indeed suggest this as, physiologically, the cerebellum responds to a learned stimulus by eliciting a pause (Jirenhed et al., 2007). This result also inevitably meant that the prediction that Crus I and Crus II would exhibit a decrease in activation as a function of long-term learning (*Repeating* < *Non-repeating*) was also not confirmed (as discussed further later).

A second result that was not in line with predictions was the absence of activation in the inferior frontal gyrus (BA 44; pars opercularis) across the temporal delay. The co-activation of the premotor cortex (BA 6) and pOp (BA 44) has been suggested to reflect the construction and maintenance of articulatory representations for subsequent action (Hickok & Poeppel, 2004, 2007; Rauschecker & Scott, 2009). Indeed, many of the neuroimaging investigations of verbal short-term memory reviewed in Chapter 1 (Section 1.5) observed evidence of such co-activation during

covert rehearsal (Awh et al., 1996; Fiez et al., 1996; Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1993; Smith et al., 1996). Pars opercularis (pOp) was not significantly active at FWE-corrected level during the temporal delay (*Temporal delay [Go] > Temporal delay [No-Go])* but was so at a lower threshold ($p < .01$). Despite the fact that the region was only significant at a lower threshold, an activation cluster in the same set of co-ordinates was observed in Experiment 2 with an identical design (but with visual sequences) at a corrected threshold (Chapter 4, Section 4.3.2.1). As such, this makes it less likely that the activation of the region at a lower threshold in the present study was a chance result. Co-activation of BA 44 with the premotor and superior temporal cortices was also not observed in the present experiment at either an FWE-corrected or lower threshold ($p < .01$) during presentation (*Presentation > Temporal delay* [*Go*]). This is at odds with research demonstrating activation of this region when sounds need to be segmented (Burton et al., 2000; Hsieh et al., 2001). Although items are presented in a temporally separated fashion and auditory sequential streaming is likely to take place, segmentation is still required insofar the perceptual-motor account proffers that the articulatory process binds unrelated items together and generates new sequential information for the purpose of retention based on coarticulation (cf. Chapter 1, Section 1.4 and Macken et al., 2016)

The discussion up to this point has centred predominantly on cerebro-cerebellar regions involved in articulatory planning. However, establishing which regions were activated as a function of perceptual processing and perhaps perceptual-motor mapping was also central to the case that there is no modality-independent phonological store. For this reason, the *Presentation > Temporal delay (Go)* contrast was hypothesised to reveal activation that could be explained by recourse to regions to specifically involved in perceptual or perceptual-motor mapping processes, over

and above the motor planning processes that were expected to be engaged throughout both phases. Consistent with the hypotheses, the results (Table 5) indicated that presentation was associated with bilateral activations along the transverse temporal gyri in the auditory cortex (BA 41) and the left planum temporale (PT) in the Spt region (BA 22). The activation of the primary auditory cortex (BA 41) is consistent with widely accepted research that bilateral activation is typically observed in the dorsal superior temporal cortices during early auditory processing of both speech and nonspeech sounds (Giraud et al., 2004; Hickok et al., 2003; Hickok & Poeppel, 2004; Meyer et al., 2005).

In particular, a left lateralization for speech has been suggested at the level of distinguishing rapidly changing temporal features (Hesling et al., 2005; Poeppel, 2003). Activation of the left planum temporale (PT) is consistent with this view as the PT is commonly activated by auditory inputs, covert articulation (Hickok et al., 2003, 2009) and studies of auditory-verbal short-term memory (Buchsbaum & D'Esposito, 2009; Koelsch et al., 2009; McGettigan et al., 2011). It is important to note that the PT does not exhibit the functional characteristics of a phonological store as it has been shown to be active during encoding and rehearsal of musical melodies as well as speech (Hickok et al., 2003) indicating that it is not specialised for verbal information. MNI co-ordinates -57, -21, 0 in the results of the *Presentation > Temporal delay (Go)* contrast are likely in the left PT and were in close proximity to those observed by McGettigan et al. (2011) during presentation of auditory stimuli (MNI: -54, -18, 6) and rehearsal (MNI: -51, -28, 3) during a high vs. low syllable load comparison in a delayed pseudoword repetition task. Simultaneous with PT activation, they also observed activity in left premotor cortex (MNI: -51, -6, 48) close to a premotor cortex cluster observed in the *Presentation | Temporal delay* conjunction (MNI: -57, 0, 42)

in the present experiment. The combination of activation in the PT and premotor cortex suggests that both regions exhibit similar response properties (Buchsbaum et al. 2001; Hickok et al., 2003), as it has been shown that the PT contains both sensory- and -motor weighted classes of cell types (Hickok et al., 2009). Critically, the peak co-ordinate clusters observed in the premotor cortex were not observed in results from the *Presentation | Temporal delay* conjunction and suggest that regions of the premotor cortex are attuned to sensorimotor integration. This may offer an explanation as to why the IFG was not observed in results of this contrast. The generation of articulatory codes and perceptual-motor mapping may have relied more on the operation of the PT (as an auditory-motor interface) and specific areas of the premotor cortex.

However, activation of the PT was only observed during presentation of sequences in the present experiment and differs from the results of previous studies that showed activation during both the presentation of auditory-verbal stimuli and their active rehearsal (Buchsbaum et al., 2001; Hickok et al., 2003). In particular, Hickok et al. (2003) observed activation of the Spt region during the presentation of nonsense sentences and piano melodies, as well as the active rehearsal of the sentences and humming of the melodies. Greater auditory-perceptual and articulatory requirements for nonsense sentences in comparison to the sequences used in the current experiment may offer some explanation as to why Hickok et al. (2003) observed activation of the PT during a rehearsal phase. There has been some consideration that the PT may be the location of a phonological store as Buchsbaum et al. (2005a) observed activation in the region during presentation and delay period phases of tasks requiring silent reading and speech perception. However, activation of the PT was only observed during auditory-verbal sequence presentation in the current experiment and not in results of the same contrast using visual-verbal sequences, nor during rehearsal of

visual-verbal sequences. Taken together with evidence that the PT is also involved in the perception and results of musical melodies (Hickok et al., 2003) suggests that the region is not specific to verbal processing but are in line with notions that it functions as an auditory-motor interface.

Many previous neuroimaging studies claimed to have localised the phonological store in various areas within the inferior parietal lobe. However, no activation of BA 40—the region most frequently identified as the location of the phonological store—was observed during the rehearsal of the sequence following sequence-presentation (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]) when, according to the phonological store theory, representations within the store would have been subject to continual retrieval and re-activation. Instead, activation along the intraparietal sulcus (IPS) was observed, specifically cytoarchitectonic subdivisions hIP3 (anterior part of the medial wall in the IPS) and hIP6 (lateral wall of the IPS) (Scheperjans et al., 2005, 2007). Studies of the subdivisions of the IPS have shown that anterior IPS maintains connections to prefrontal regions and posterior IPS is connected to early and higher visual regions (Bray et al., 2013; Mars et al., 2011; Uddin et al., 2010). Area hIP6 has been shown to be involved in a variety of functions such as action inhibition, covert word generation and delayed match to sample (Richter et al., 2019).

Activations in the inferior parietal lobe in early neuroimaging studies of short-term memory were often not discussed outside of a phonological store framework, yet there is good evidence suggesting that subdivisions of the IPS as well as other regions across the parietal lobe are involved in a multitude of functions. The results of the *Presentation > Temporal delay (Go)* contrast clearly provided evidence of modality-specific processing in tandem with motor processing. Phonological store theory would

state that alongside such processing, a region that holds modality-independent representations separate to motor and perceptual regions should be observed—however no such region was found to be consistent with this notion. Finally, the neural correlate of the phonological store should have been evident across both task phases and yet no parietal region was consistently active across presentation and the temporal delay in the present experiment, contradicting the notion of a phonological store localised to anywhere across the parietal lobe (Awh et al., 1996; Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1993; Smith et al., 1996, see also Chein & Fiez, 2001). The fact that this was not the case also indicates that the parietal activations that were observed during the temporal delay could not have reflected the operation of a phonological store.

A second key aim of this experiment was to test the hypothesis that the long-term learning of a verbal sequence—as witnessed in the form of the Hebb repetition effect (Hebb, 1961)—is based on the increasing fluency of the vocal-motor plan generated to retain and reproduce that sequence over the short term. While the behavioural data from Experiment 1B replicated the Hebb effect, the predicted decrease in cerebellar activity across the instances of the repeating sequence—simultaneous with the increase in behavioural recall accuracy—was not observed. The images and parameter estimate graphs in Figure 13 indicate that repeating 'Hebb' and non-repeating 'Filler' sequences were processed almost identically across the cerebellum and changes in excitability were not observed over cycles. Lobules HVI and HVIIB active during the *Temporal delay (Go) > Temporal delay (No-Go)* contrast were also significantly active when comparing repeating and non-repeating sequences in Experiment 1B but again the parameter estimates showed that it was the difference between activity modelled in the polynomial regressors (linear and quadratic trends)

compared with activity in the non-modulated regressors ('Hebb' and 'Filler', where polynomial expansions were not used) that actually gave rise to these significant effects in lobules HVI and HVIIB.

It is not entirely clear at this point as to why observing a decrease of the BOLD-signal in the cerebellum simultaneously with an increase in behavioural accuracy was not observed. One possibility, however, is that given that the present experiment involved *auditory*-verbal sequences, the experiment was not optimal for testing the prediction that repetition of a verbal-motor plan would result in long-term learning. Although the hypothesised learning effects were not observed in this experiment using auditory-verbal sequences, it cannot be guaranteed that cerebellar learning of verbal sequences does not occur at all, as evidence of motor-skill learning in the cerebellum is widely acknowledged (Doyon et al. 2002; Imamizu et al. 2000; Jueptner et al., 1997; Penhune & Doyon 2005; Ramnani & Passingham, 2001; Sakai et al., 2002b). For example, in comparison to auditorily presented sequences, long-term learning of visually presented sequences has been shown to rely more on articulatory planning, as evidenced by a greater effect of articulatory suppression on Hebb sequence learning with visual- compared to auditory-verbal sequences (Sjöblom & Hughes, 2020). The authors posited that visual-verbal sequences do not benefit from passive acoustic-based perceptual organisation-based learning as auditory sequences do and hence the learning of visual-verbal sequences may rely more on articulatory planning. One aim of Experiment 2 (cf. Chapter 4), then, will be to examine whether the excitability decreases predicted in Experiment 1B will indeed be observed with visual-verbal stimuli.

3.5 Conclusions

In summary, the present experiment supports the argument that verbal serial short-term memory performance can be explained largely in terms of the recruitment of the motor planning system. Activation of prefrontal and premotor cortices during the temporal delay was in line with this view, as was the activation of premotor projecting lobules HVI and HVIIB of the cerebellum. The consistency of premotor cortex activation across presentation and temporal delay phases suggests that articulatory rehearsal was the common process to both task phases. Critically, no region in the inferior parietal cortex was found to be active across both task phases, thereby contradicting predictions derived from the phonological store theory. Activation of specific regions within the premotor cortex—that were not observed in results of any other contrast—simultaneous with evidence of modality-specific processing across the temporal lobe indicated that a perceptual-motor mapping process may occur. This result further supports the perceptual-motor approach to verbal serial short-term memory and contradicts claims of phonological store theory that auditory information does not require such mapping and gains obligatory access to a modality-independent store.

The present experiment involved only *auditory*-verbal sequences. In Experiment 2 (Chapter 4), the design of Experiment 1 was replicated but, crucially, using *visual*-verbal sequences. This should further adjudicate between the two theoretical approaches of interest in the present thesis (phonological store theory and perceptual-motor account): If verbal serial short-term memory is indeed a by-product of motor planning, then any commonalities due to the involvement of articulatory planning necessary for recall, regardless of presentation modality, will be reflected in activation across the prefrontal and premotor cortices, along with the cerebellum.

Furthermore, activation during sequence presentation in visual-perceptual and motor planning regions— that are not observed during any other task phases—may indicate the operation of a perceptual-motor mapping process. If activation in such regions is observed whilst no one region in the inferior parietal cortex is identified across experiments, results would refute a central claim of the classical model—that items become modality independent, devoid of any perceptual or motoric features upon entering a proposed phonological store (Baddeley & Hitch, 1994). As noted, the experiment will also examine again the basis of verbal sequence learning given that visual-verbal sequences are predicted rely more heavily on motor planning processes.

CHAPTER IV: VISUAL-VERBAL SERIAL SHORT-TERM MEMORY AND LEARNING

4.1 Introduction

Verbal serial short-term memory and learning phenomena have typically been explained by recourse to the existence of a passive phonological store supported by the operation of an active articulatory rehearsal process (Baddeley et al., 1984; Salamé & Baddeley, 1982; Vallar & Baddeley, 1984). The articulatory rehearsal process has been mapped to the premotor cortex, inferior frontal gyrus and cerebellar lobules HVI and HVIIA (Crus I) (e.g., Chen & Desmond, 2005a,b Desmond et al., 1997). But most cognitive neuroscientific research on verbal short-term memory has been dominated by attempts to isolate activation indicative of a phonological store. While varying locations across the inferior parietal cortex have been suggested as the home of the store (Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996), closer inspection indicates that the interpretation of results was often incompatible with the phonological store theory upon which they were based (cf. Chapter 1, Section 1.5).

Rather than posit the existence of a specialised phonological store, the perceptual-motor account reconceptualises performance over the short term by recourse to articulatory planning and perceptual processing (Hughes & Marsh, 2017; Hughes et al., 2009, 2016; Jones et al., 2006, 2004; Macken et al., 2016; Sjöblom & Hughes, 2020). The results of Experiment 1 (Chapter 3) lent clear support to the perceptual-motor account: Significant activation of the prefrontal cortex (area 46), premotor cortex (BA 6), and cerebellar lobules IV, HVI and HVIIB was across a temporal delay between sequence presentation and recall, suggesting that an

articulatory rehearsal process was used to establish and retain the sequence prior to recall (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]). The activation of subdivisions within the premotor cortex during both presentation and the temporal delay—whilst no activations in the parietal cortex, could be taken to reflect the existence of a phonological store (cf. Chein & Fiez, 2001)—challenge the competing store-based account and centre the articulatory rehearsal process for verbal serial short-term performance. Instead, activation across the temporal delay in subdivisions of the intraparietal sulcus (IPS) were explained with reference to their connections to the prefrontal cortex and early visual regions (Bray et al., 2013; Mars et al., 2011; Uddin et al., 2010). The subdivisions of the IPS have been shown to facilitate variety of functions such as action inhibition, covert word generation and delayed match to sample (Richter et al., 2019) and do not exhibit the characteristics of a specialised phonological store. Moreover, modality-specific activation during presentation of auditory sequences was observed in the primary auditory cortex and planum temporale. When taken together, results across several contrasts from Experiment 1 can be explained in terms of a network of regions involved in the assembly and cyclic iteration of an articulatory plan.

The results of Experiment 1B demonstrated that excitability decreases in prefrontal and premotor projecting lobules in the cerebellum were associated with both repeating and non-repeating sequences. It is not entirely clear as to why a decrease in the BOLD-signal was not observed exclusively for repeating sequences however the possibility remains that auditory-verbal sequences were not optimal for testing that prediction. Behavioural evidence has demonstrated differences in long-term learning of auditory and visual sequences: visual-verbal sequence learning relies more heavily on articulatory planning than auditory-verbal sequence learning (Sjöblom & Hughes,

2020) which could mean that an excitability decrease in cerebellum (indicative of motor skill learning) is more likely to be observed using visual-verbal sequences in Experiment 2B. The study of the neural basis of verbal serial short-term memory with visual sequences and its (qualitative) comparison with that using auditory sequences in Experiment 1 has the potential to adjudicate further between the perceptual-motor account and the phonological store theory.

Both accounts would predict differences in activations during the presentation phase of the current experiment compared to the presentation phase of Experiment 1 due to the different brain regions implicated in the encoding of visual compared to auditory stimuli. However, the key difference between the accounts is that on the phonological store theory there should be a common set of regions active across the two experiments that reflect the action of a modality-independent phonological store. Moreover, these activations should be found in regions that are unlikely to play a role in articulatory processes because the store is said to be independent from such processes (Baddeley, 2007). Indeed, not only should it be a region that is not involved in articulatory processes, but it should be a region that is not involved in any function other than phonological storage (Buchsbaum & D'Esposito, 2008). The defining characteristic of the *phonological* store is that it holds abstract, modality-independent, representations of verbal input regardless of the sensory modality by which that input was registered. Thus, representations in the phonological store lack any perceptual (or motoric) features. The only difference between auditory and visual presentation on the phonological store theory is that auditory input has been argued to gain automatic access to the phonological store whereas visual input requires the articulatory rehearsal system to convert the input into phonological form (e.g., Baddeley, 1986; see Chapter 1, Figure 1). The phonological store account would posit that activation of auditory-

perceptual and motor regions would be observed alongside a phonological store during sequence presentation, but that no need (or region) would exist for the mapping of auditory input onto motor processes.

On the perceptual-motor account, similarities in activations across the two experiments should indeed be readily attributable to the action of motor processes or the process of mapping perceptual products onto such processes. This follows from one of the most fundamental difference between the perceptual-motor account and the phonological store theory: Instead of refreshing decay-prone items held within a store or converting visually presented verbal items into a phonological code, articulatory rehearsal is the very means by which the to-be-remembered items are bound into a coherent, reproducible, sequence (e.g., Hughes & Jones, 2009; Jones et al., 2006, 2004). As such, there should be evidence of the action of such articulatory processes not only across different task-phases (cf. Experiment 1) but also across input-modalities. Results from Experiment 1A—and predictions of the perceptual-motor account—are consistent with the notion that perceptual-motor mapping is required during verbal serial short-term memory tasks. Activation of the planum temporale (*Presentation* > *Temporal delay* [*Go*]) is in line with suggestions that the region functions as an auditory-motor interface and may have been involved in mapping auditory input to a motor plan (Buchsbaum & D'Esposito, 2008). The perceptual-motor account proffers that *both* auditory and visual information are mapped to a motor plan that is co-opted for serial rehearsal given the inherent temporal and sequential features of articulation. Motor planning regions are therefore proposed to interact with modality-specific perceptual processing regions or those involved in sensorimotor integration.

The current study therefore extends on Experiment 1 by presenting the to-be-remembered items visually. A qualitative comparison across the experiments should adjudicate further between the perceptual-motor and phonological-store based accounts. The activation of motor planning regions in both experiments, combined with distinct auditory and visual operations—whilst no one region consistent with the notion of a phonological store is observed across experiments—will further support the notion that motor planning and perceptual processes are sufficient in accounting for verbal serial short-term memory performance. It is hypothesised that regions known to be involved in motor planning and shown to be active in Experiment 1 will also be observed during the retention of visual-verbal sequences: Activation of the premotor cortex (BA 6) and its cerebellar targets, lobules HIV-HVI, HVIIB and HVIII (Kelly & Strick, 2003; Lu et al., 2007) are predicted to be observed across the temporal delay. If such activation is observed, results would be interpreted to reflect feedforward error-predictions necessary for articulatory motor planning (i.e., covert articulation). Activation of the prefrontal cortex (area 46) and its cerebellar target lobule HVIIA (Crus I and Crus II) should be active on the basis that this cortical area is involved in maintaining higher-order representations across delays (Goldman-Rakic,1991) necessary for the achievement of action-related goals (Passingham, 1996). However, the direction of activation in lobule HVIIA observed in Experiment 1A (*Temporal delay [Go] > Temporal delay [No-Go*]) was not as predicted. This finding therefore complicates current predictions. On the one hand, assessing whether the originally hypothesised direction of activation is observed in the present study may provide insight into whether lobule HVIIA contains forward models that predict the responses of the premotor cortex—to commands issued by the prefrontal cortex—regarding higher-order rule-based information (Ramnani, 2006). On the other hand,

if the same pattern of activation is observed in results of the *Temporal delay* (*Go*) > *Temporal delay* (*No-Go*) contrast across modalities, then results will warrant a reassessment of lobule HVIIA contribution to articulatory planning and may suggest that the acquisition of forward models occurs on a different timescale to originally predicted.

Activation during presentation is predicted to reveal areas across the occipitotemporal cortex indicative of visual-verbal perceptual processing and regions within the premotor cortex over and above those predicted to be consistently active during presentation and the temporal delay. If such regions are observed concurrent with occipitotemporal activations—as was the case during auditory presentation (auditory cortex, premotor cortex and planum temporale)—results will suggest that a perceptual-motor mapping process also occurs in the visual domain. These results could demonstrate that visual objects are mapped onto motor acts via an occipitotemporal pathway (Goodale & Milner, 2018, 1992; Kuypers et al., 1965; Turner et al., 1980). Results of a conjunction analysis are predicted to show the same activations in the prefrontal and premotor cortices similar as those observed in results of the auditory conjunction. At the same time, no region consistent with a phonological store – functionally and anatomically separate from perceptual and motor regions— will be evident. When taken together, results from contrasts across both experiments are expected to be interpretable mainly by recourse to the involvement of articulatory planning and perceptual processing to verbal serial short-term memory performance, obviating the need for a phonological store.

The present experiment also consisted of an Experiment 2B that replicated Experiment 1B but again with visual rather than auditory stimuli. A decrease in the BOLD-signal across cerebellar lobules active during rehearsal in Experiment 2

(*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]) is hypothesised to be observed as a result of long-term learning of a repeating Hebb sequence (Hebb, 1961).

4.2 Methods

4.2.1 Subjects

Twenty-seven subjects were recruited from Royal Holloway, University of London's Experiment Management System. Subjects were remunerated £10 and received a structural image of their brain following data analysis. The experiment was terminated early for four subjects who felt discomfort in the MRI scanner. Their data was not included. Following fMRI data quality assurance, two of the remaining twenty-three subjects were excluded from the functional data set due to excessive movement (see Chapter 2, Section 2.7.2). The final number of subjects included in data analysis was twenty-one (Mean age = 23 years, SD = 5 years, female = 12, male = 9).

4.2.2 Stimulus Design

All aspects of the method were the same as for Experiment 1 except for the to-be-remembered stimuli. The same letters as presented in Experiment 1 were now presented visually (to re-cap, these were the phonologically similar letters 'B' 'C' 'D' 'T' 'P' 'V' 'G' and the phonologically dissimilar letters 'F' 'K' 'L' 'R' 'Y' 'H' 'Q') using Psychtoolbox and MATLAB. They were presented individually for 700 ms with no interstimulus interval. Dimensions of the stimuli on the screen were programmed using the Psychtoolbox function 'TextSize' and set to 40. The default text style of Windows operating systems was the font 'Segoe.' Stimuli were presented in the centre of the projector screen in the MRI unit using Psychtoolbox (Chapter 2, Section 2.6.1).

4.3 Results

4.3.1 Behavioural Results

4.3.1.1 Experiment 2A (Independent test of engagement in articulatory rehearsal/phonological storage).

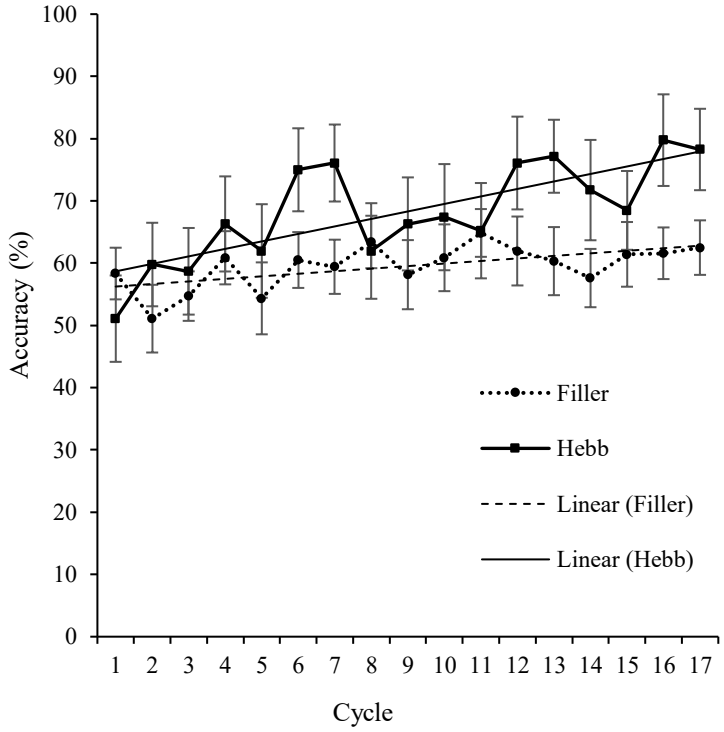
Recall of phonologically similar items (M = 57.1%, SE = 3.18%) was significantly poorer than recall of dissimilar items (M = 62.1%, SE = 3.05%), $t(22) = 1.96$, p =.03 (one-tailed), thus replicating the phonological similarity effect. Again, the replication of the effect confirms, from a perceptual-motor perspective that subjects engaged in articulatory planning and, from the perspective of the phonological store theory, that the phonological store was used.

4.3.1.2 Experiment 2B (Hebb Sequence Learning).

Figure 14 shows the percentage of items correctly recalled on Hebb trials and on Filler trials as a function of cycle. Performance for Filler trials was calculated in the same way as for Experiment 1B: Filler performance was calculated as the average recall across the two or three Filler trials presented at a given cycle. As shown in Figure 14, recall of the Hebb sequence (M = 67.3%, SE = 4.48) was better than that of Filler sequences (M = 59.5%, SE = 3.32). An overall increase in the recall of the Hebb sequence is also observable across the 17 cycles, ranging from M = 51% in cycle 1 to M = 78% in cycle 17, whereas little such improvement is seen across cycles for Filler trials. A 2 (List-type) × 17 (Cycle) Repeated-Measures ANOVA revealed a main effect of List-type, $F(1, 22) = 9.216$, $MSE = 1294.3$, $p < .006$, $\eta_p^2 = .295$, reflecting the enhanced recall of Hebb sequences in comparison to Filler sequences generally (i.e., a replication of the Hebb effect). There was also a main effect of Cycle, $F(16, 352) = 1.776$, $MSE = 597.55$, $p < .033$, $\eta_p^2 = .075$, but the List-type by Cycle interaction was not significant on this occasion, $F(16, 352) = .848$, $MSE = 683.98$, $p < .630$, $\eta_p^2 = .037$.

Figure 14

Percentage recall accuracy for Filler and Hebb sequences as a function of cycle, including linear trendlines, in Experiment 2B.



4.3.2 fMRI results

4.3.2.1 Experiment 2: Temporal delay (Go) > Temporal delay (No-Go).

As predicted, activations in premotor cortex BA 6 (Figure 15B) in the precentral gyrus, and inferior frontal gyrus, (pars opercularis; BA 44 , Figure 15C) were observed during the temporal delay when participants were expected to have continued engaging in articulatory planning (Go trials) compared to when there was no reason for them to do so (No-Go trials). While no activation was observed in BA 46 in the middle frontal gyrus, the co-ordinates of BA 9 (MNI: -42, 30, 27) likely fall within area 9/46 (Petrides & Pandya, 1999) and are shown in the activation map (Figure 15A). Cerebellar lobules HVI and HVIIB were also significantly active and the parameter estimate plots (Figures 16A & 16C) show similar patterns of activation to those in the frontal cortex.

These results, other than the activation of BA 44, replicate those observed in the equivalent contrast in Experiment 1A. For comparative purposes, activation observed in results from this contrast in the auditory domain are presented alongside those from the present (visual) results in Figures 17A-C. The parameter estimate plot for Crus II (Figure 16B) does not demonstrate increased activation above baseline as originally hypothesised but replicates the activity observed in Crus II with auditorily presented sequences in Experiment 1A. Figure 16B shows instead a decrease in activation. In addition, whilst activation of vermal lobule IV was observed during delay-period rehearsal of auditory sequences (Experiment 1A), it was not active during delay-period rehearsal of visual sequences.

Figure 15

Parameter estimate and sagittal slices of MNI152 template image showing activation from results of Temporal delay (Go) > Temporal delay (No-Go) {T} contrast in Experiment 2. A: middle frontal gyrus area 9/46 (-42, 30, 27). B: pre-central gyrus BA 6 (MNI: -48, -3, 45). C: inferior frontal gyrus (pars opercularis; BA 44) (crosshair located in [MNI: -51, 12, 3], parameter estimate plot in [MNI: 33, 24, -3]).

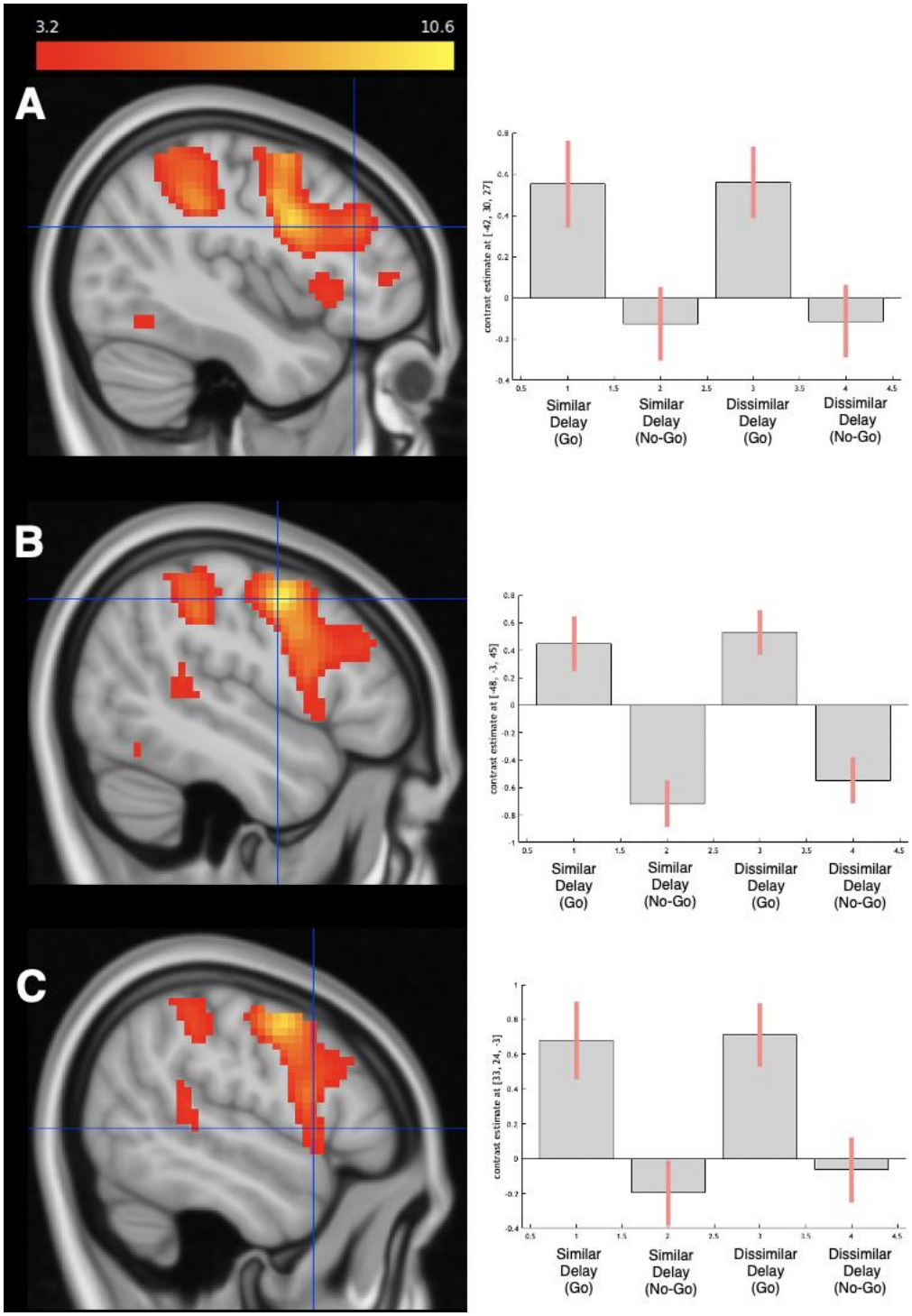


Figure 16

Parameter estimates and images from the results of the Temporal delay (Go) > Temporal delay (No-Go) {T} contrast in Experiment 2. A: Coronal slice of MNI152 template showing activation in cerebellar lobule VI and parameter estimate in (33, -57, 27). B: Sagittal slice of MNI152 template showing activation in lobule HVIIA (Crus II) and parameter estimate (9 -78, -33). C: Coronal slice of MNI152 template showing activation in cerebellar lobule HVIIB and parameter estimate in (27, -69, -48).

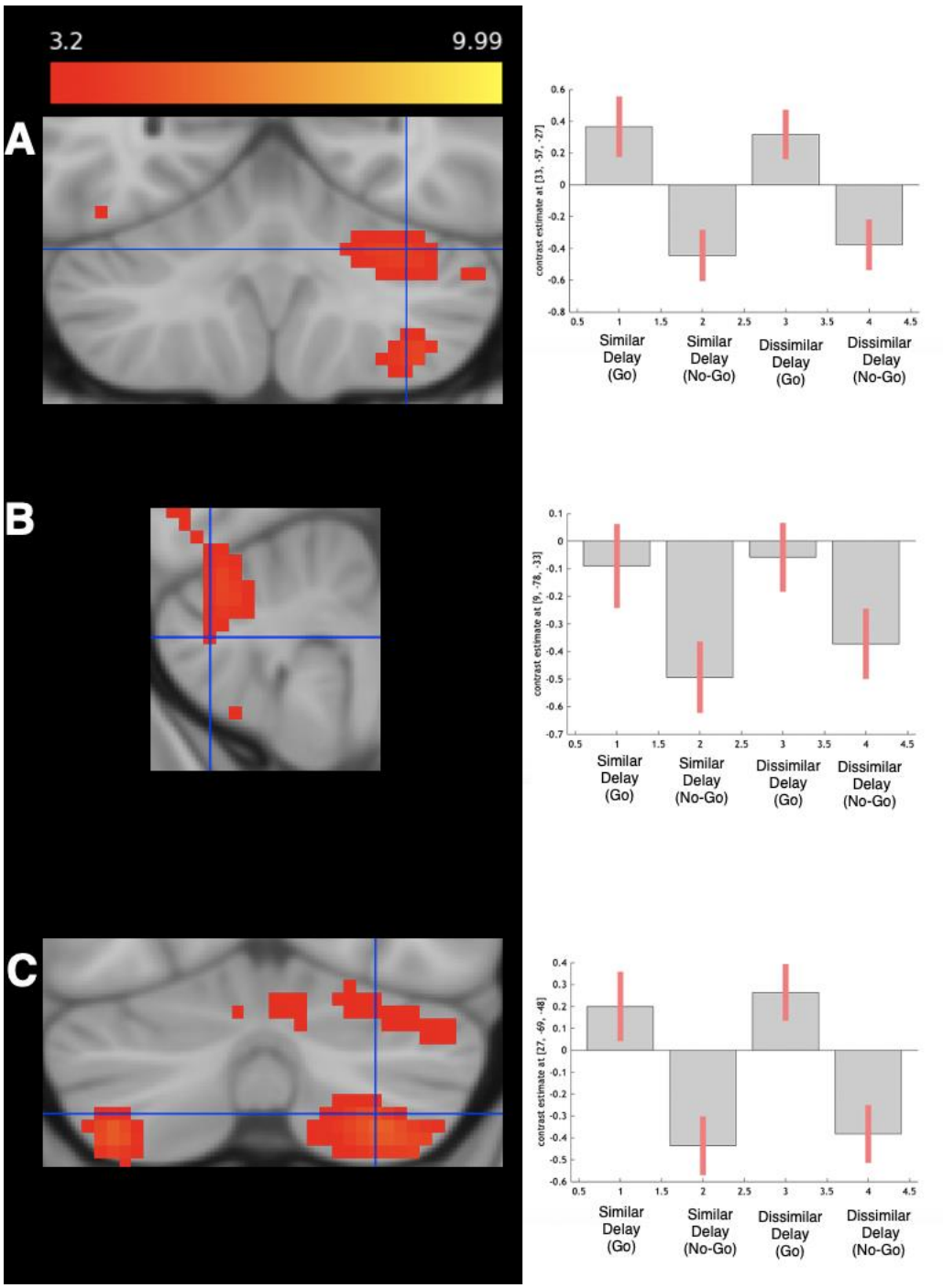


Figure 17

Comparison images across Experiments 1 and 2. A: Sagittal slices of MNI152 template image showing activation from results of the Temporal delay (Go) > Temporal delay (No-Go) {T} contrast in the premotor cortex. B: Sagittal slices of MNI152 template image showing activation from results of the Temporal delay (Go) > Temporal delay (No-Go) {T} contrast in the prefrontal cortex, premotor cortex and parietal lobe. C: Coronal slice of MNI152 template showing activation from results of the Temporal delay (Go) > Temporal delay (No-Go) {T} contrast in cerebellar lobules HVI, HVIIA and HVIIB.

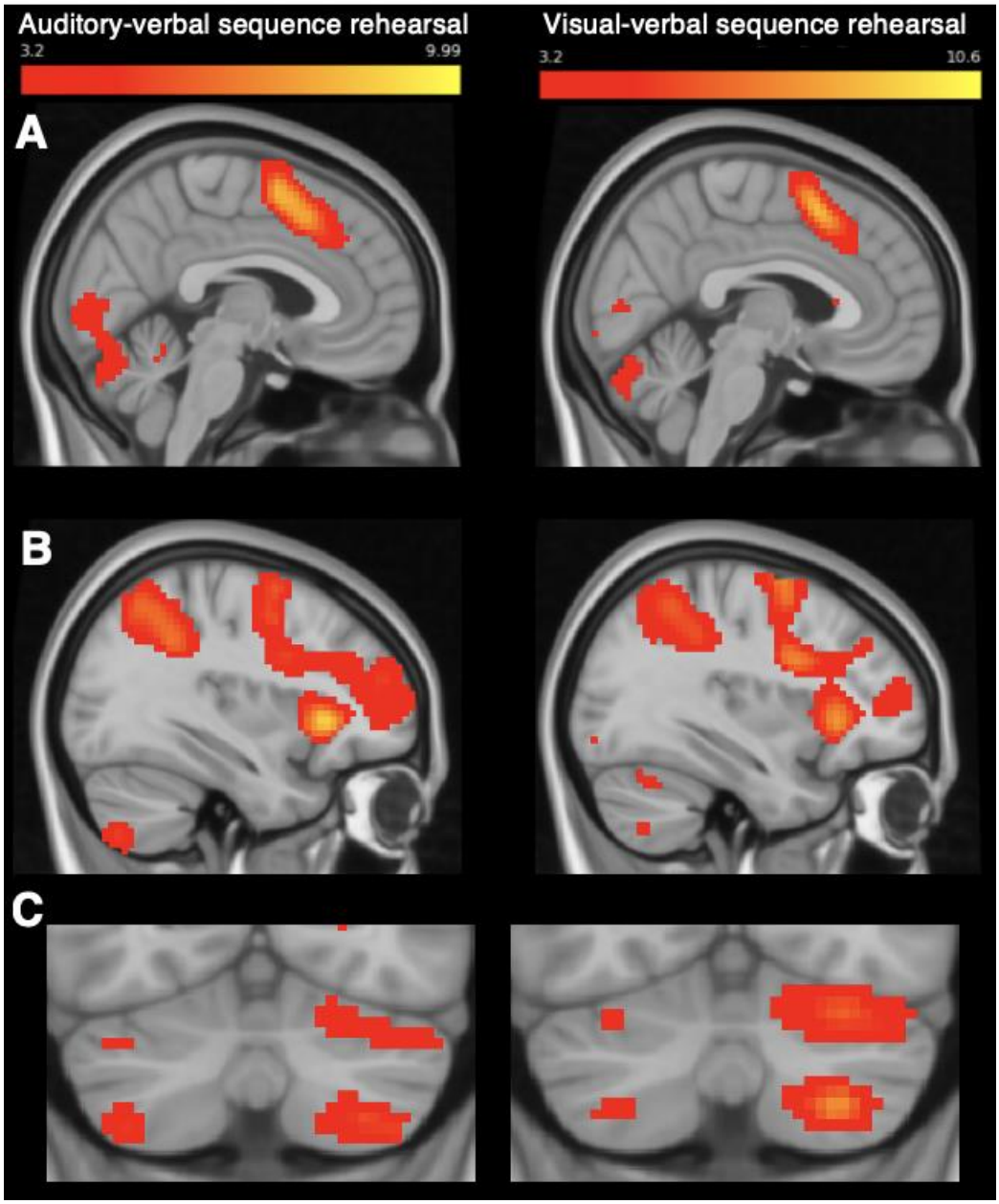


Table 8

*Results of a paired t-test, Temporal delay (Go) > Temporal delay (No-Go) {T} contrast in Experiment 2 (FWE-corrected for multiple comparisons, p<.05; *Prefrontal and Premotor cortices SVC; **Whole Cerebellum SVC; † Prefrontal, & Premotor cortices SVC and FWE-corrected for multiple comparisons, p< .05; ‡ Whole Cerebellum SVC and FWE-corrected for multiple comparisons, p <.05).*

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Frontal Lobe							
Prefrontal cortex							
Superior frontal gyrus	6.67	5.93	33	6	60	8	
Inferior frontal gyrus pars opercularis†	6.17	5.57	-51	9	18	44	54.20%
Middle frontal gyrus*	6.01	5.44	-42	30	27	9	
Middle frontal gyrus†	5.81	5.29	42	36	30	9	
Middle frontal gyrus†	5.24	4.84	36	51	9	10	
Inferior frontal gyrus pars opercularis*	5.2	4.81	-54	12	6	44	66.50%
Medial orbitofrontal cortex*	4.93	4.59	21	48	-12	11	
Medial orbitofrontal cortex	4.93	4.59	21	48	-12	11	
Middle frontal gyrus*	4.13	3.92	-36	48	6	10	
Middle frontal gyrus*	4.05	3.85	-33	51	15	10	
Medial orbital gyrus*	3.78	3.62	-24	48	-12	11	
Premotor Cortex							
Pre-central gyrus†	10.5	Inf	-48	-3	45	6	
Pre-central gyrus†	10.04	Inf	-3	12	51	6	
Pre-central gyrus†	9.58	7.79	-42	3	30	6	
Pre-central gyrus*	6.67	5.93	33	6	60	6	
Pre-central gyrus*	4.19	3.97	42	6	33	6	
Superior frontal gyrus*	4	3.81	51	9	42	6	
Superior frontal gyrus*	3.94	3.75	51	3	48	6	
Pre-central sulcus*	3.22	3.12	30	-6	48	6	
Insular Cortex							
Insular cortex†	8.4	7.09	-27	27	0		
Insular cortex†	6.93	6.12	33	24	-3	Id7	51.1%
Parietal Lobe							
Intraparietal sulcus	7.07	6.21	-24	-63	42	7	
Intraparietal sulcus	6.96	6.14	-42	-39	39	7	
Intraparietal sulcus	6.47	5.79	39	-42	39	7	
Intraparietal sulcus	6.36	5.7	30	-54	48	7	
Intraparietal sulcus	6.05	5.47	-36	-51	54	7	
Parieto-Occipital Fissure	4.84	4.52	15	-69	48	7	
Cerebellum							
Lobule HVIIB**	7.59	6.57	27	-69	-48		
Lobule HVI**	7.39	6.43	33	-57	-27		

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Lobule HVI**	4.97	4.62	12	-75	-21		
Lobule HVI**	4.69	4.39	-6	-75	-24		
Crus II	4.4	4.15	9	-78	-33		
Lobule HVIIB**	3.93	3.75	-33	-66	-51		
Lobule HVI**	3.9	3.72	-24	-60	-30		

4.3.2.2 Experiment 2: Presentation > Temporal delay (Go).

The process of mapping perceptual representations onto an articulatory plan was of particular interest in this thesis. Critically, the region(s) involved in perceptual-motor mapping should not be found to be active in the results from the conjunction analysis. Whilst regions involved in perceptual processing are predicted to differ across studies (and thus modalities, cf. Chapter 3: Section 3.3.2.2), similar premotor activations would highlight the centrality of articulatory planning in verbal serial short-term memory. Results of this contrast were predicted to reveal regions across the occipito-temporal cortex and premotor cortex. As predicted, regions within the occipitotemporal and premotor cortices were observed. Peak cluster co-ordinates within the premotor cortex were not observed in results from any other contrast across both the present and auditory study. Areas of the lateral occipital cortex (hOc4lp; Malikovic et al., 2016) and the fusiform gyrus (BA 19) were shown to exhibit greater activity during presentation of the visual-verbal sequences in comparison to the temporal delay (cf. parameter estimate plot; Figure 18A). Both BA 19 and hOc4lp fall within the extrastriate cortex. Activation was also observed in the temporo-occipital region (BA 22; left posterior middle temporal gyrus). For the activations observed in the right superior temporal gyrus (MNI: 51, -42, 9) and right intraparietal sulcus (MNI: 24, – 42, 39), a group-level activation map was overlaid on each individual subject's T1 structural image to ensure specificity during localisation as it was not possible to discern the location of the activation on the MNI152 template. The gross anatomy and

Brodmann areas reported in the results table for both sets of coordinates (Table 9) were observed in 20 of the 21 subjects. A region within the premotor cortex, specifically area 6MR in the pre-SMA (Amunts et al., 2019; Amunts & Zilles, 2015), was also active during presentation. The activation of this area was also observed during the presentation of auditory-verbal sequences (cf. Chapter 3, Section 3.3.3.2). Overall, then, the results confirmed the prediction that when comparing the data from this contrast across Experiments 1 and 2, similar clusters of activation—particularly of regions associated with articulatory planning—together with differences in activation specifically during presentation would be observed. Bilateral Crus II and lobule HVIIB activation (Figures 18B and 18C) during presentation of visual sequences in the present experiment was not observed during presentation of auditory sequences.

Figure 18

Parameter estimatse and images from results of Presentation > Temporal delay (Go) {T} contrast in Experiment 2. A: Sagittal slice of MNI152 template image showing activation in visual cortex and parameter estimate in lateral occipital cortex (hOc4lp, MNI coordinates: -30, -87, -6). B: Coronal slice of MNI152 template image showing activation and parameter estimate in cerebellar lobule HVIIA (Crus II) (MNI: -30, -87, -6). C: Sagittal slice of MNI152 template image showing activation and parameter estimate in lobule HVIIB (MNI: 27, -66, -48).

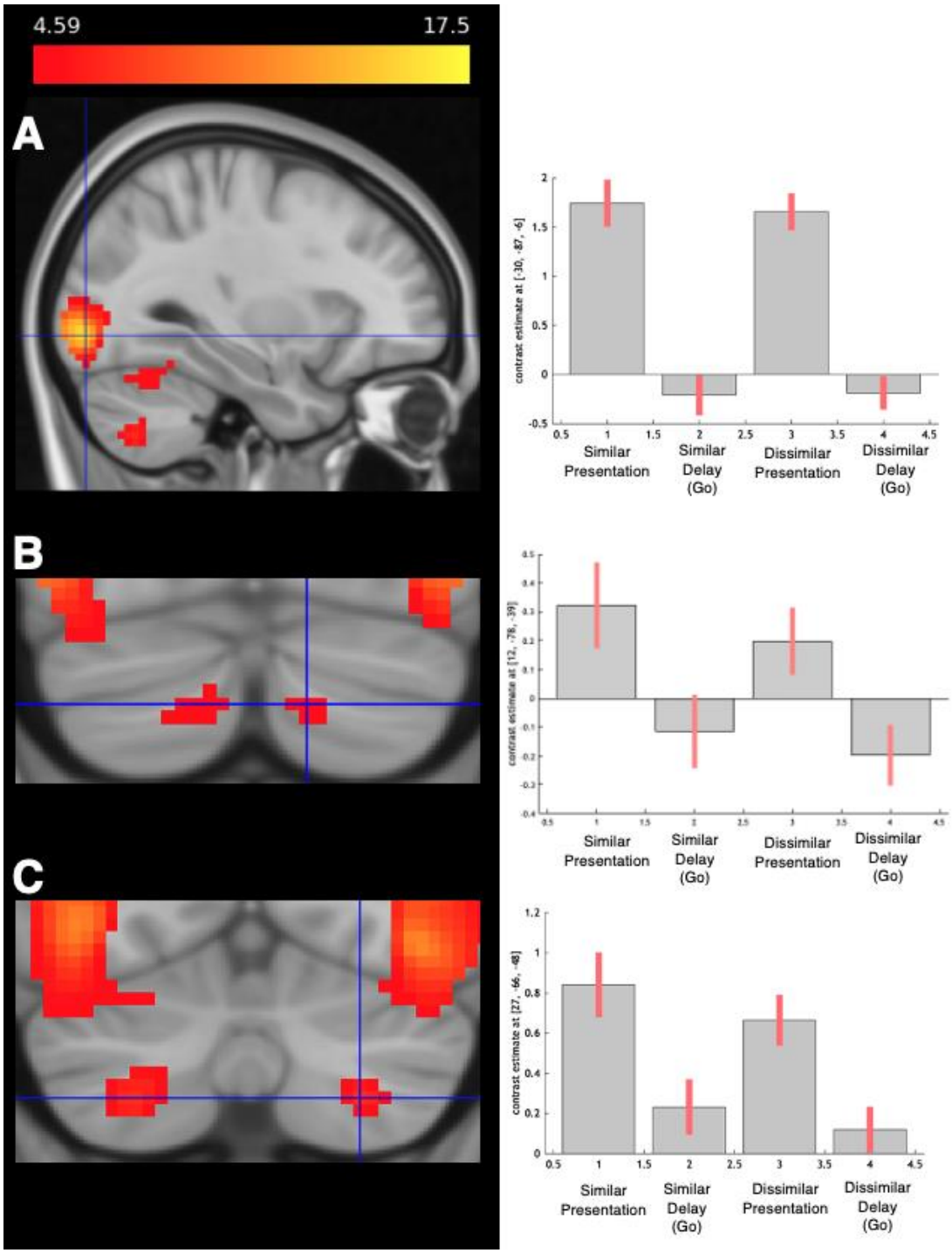


Figure 19

Comparison images across Experiments 1 and 2. A: Sagittal slice of MNI152 template image showing activation from results of Presentation > Temporal delay (Go) {T} contrast in the auditory and premotor cortices (Experiment 1, left) and visual and premotor cortices (Experiment 2, right). B: Sagittal slices of MNI152 template image showing activation from results of Presentation > Temporal delay (Go) {T} contrast in the premotor cortex (Experiment 1, left and Experiment 2, right).

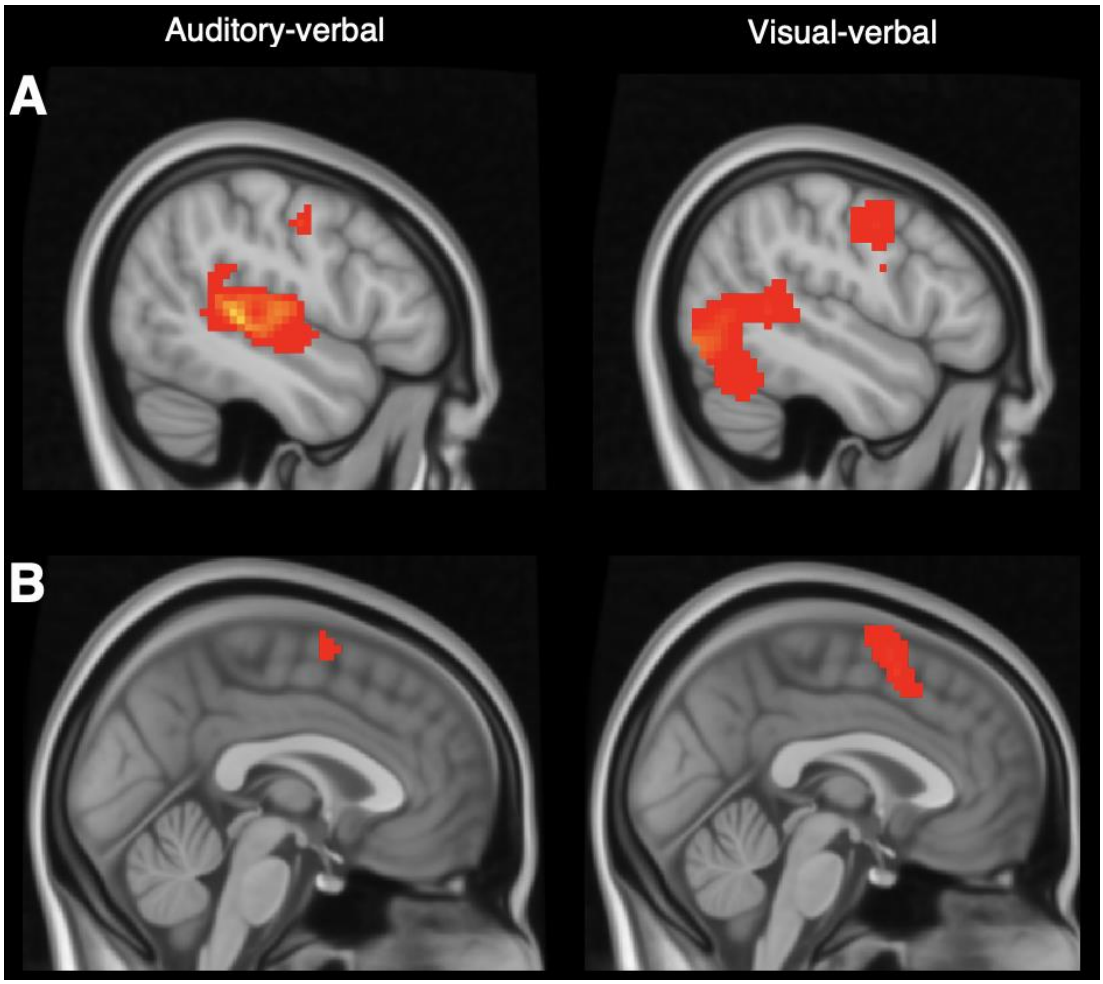


Table 9

Results of factorial analysis of variance (ANOVA), Presentation > Temporal delay (Go) {T} contrast in Experiment 2 (FWE-corrected for multiple comparisons, p<.05).

Gross Anatomy	F	Z	Coordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Frontal Lobe							
Premotor cortex							
Precentral gyrus	9.13	Inf	-6	6	60	6mr	73.7
Precentral gyrus	8.25	7.34	-45	-3	42	6	
Precentral gyrus	7.19	6.56	45	6	27	6	
Precentral gyrus	6.77	6.22	54	9	39	6	
Precentral gyrus	6.03	5.63	-39	0	30	6	
Precentral gyrus	5.49	5.18	-63	6	15	6	
Temporal Lobe							
Superior temporal gyrus	8.57	7.56	51	-42	9	22	
Middle temporal sulcus	8	7.16	-48	-48	6	22	
Basal Ganglia							
Internal Capsule anterior limb	5.58	5.26	-15	9	3		
Parietal Lobe							
Intraparietal sulcus	7.14	6.52	30	-63	30	7	
Intraparietal sulcus	5.41	5.12	24	-42	39	7	
Intraparietal sulcus	4.96	4.72	30	-51	48	7	
Occipital Lobe							
Lateral occipital cortex	17.33	Inf	30	-90	-3	hOc4lp	66.1
Lateral occipital cortex	15.44	Inf	-30	-87	-6	19	
Fusiform gyrus	11.5	Inf	-42	-72	-6	19	
Fusiform gyrus	11.37	Inf	42	-63	-12	19	
Cerebellum							
Lobule VIIb	7.2	6.56	-27	-69	-48		
Lobule VIIb	6.94	6.36	27	-66	-48		
Crus II	5.7	5.36	-12	-75	-42		
Crus II	5.46	5.16	12	-78	-39		

4.3.2.3 Experiment 2: Temporal delay (Go) > Presentation.

For the sake of completeness, this contrast was conducted to assess whether any regions were significantly active during rehearsal (in the absence of any stimuli) above and beyond activation during the presentation phase. Areas in the cuneus (BA 17; BA 19 and hOc3d, dorsal V3; Kujovic et al., 2013) were significantly active across the temporal delay compared to presentation. The cluster of activation within BA 19 in results of this contrast (Table 10) differ in location to that observed in BA 19 during visual-verbal sequence presentation (*Presentation > Temporal delay* [*Go*]). These results suggest that subregions within BA 19 responded differently to distinct task phases in the present experiment. V1 (BA 17) is the primary visual cortex and hOc3d is a cytoarchitectonic subdivision of dorsal V3 in the extrastriate cortex.

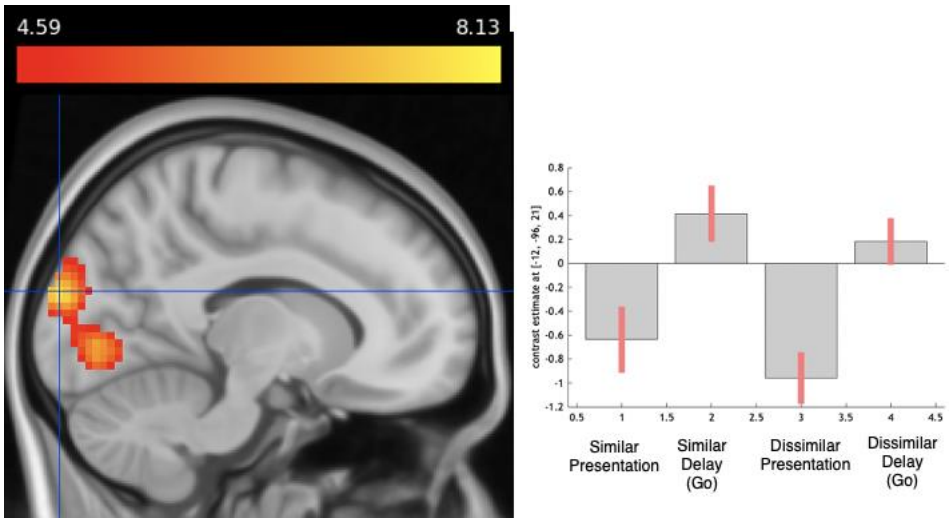
Table 10

Results of factorial analysis of variance (ANOVA), Temporal delay (Go) > Presentation {T} contrast in Experiment 2. FWE-corrected for multiple comparisons, p<.05.

Gross Anatomy	F	Z	Coordinates (x, y, z)			Brodmann Area	Probabilistic Cytoarchitecture (if available)
Occipital Lobe							
Cuneus	8.1	7.23	18	-93	21	19	
Cuneus	7.75	6.97	-12	-96	21	hOc3d [V3d]	50.3
Cuneus	6.69	6.16	-12	-81	0	17	

Figure 20

Parameter estimate and sagittal slice of MNI152 template image from results of Temporal delay > Presentation {T} contrast in Experiment 2 showing activation in the visual cortex and parameter estimate in the cuneus, area hOc3d (MNI: 12, -96, 21).



4.3.2.4 Experiment 2: Presentation | Temporal delay (Go) (Conjunction).

A conjunction analysis was conducted to assess which regions may have been active across both the presentation and temporal delay (Go) phases and whether one or a number of these could plausibly be characterised as the location of a phonological store. Based on the approach throughout this thesis, it was hypothesised that no such region(s) would be identified; instead, any regions active across the two phases would be indicative of the common engagement of articulatory planning processes. The activation of a prefrontal region in the middle frontal gyrus (BA 9) and premotor region in the precentral gyrus (BA 6) common to both task phases confirmed these predictions as similar clusters of activation in the same regions were observed in the results of the same contrast in Experiment 1 (for comparison of premotor activation see Figure 22). In results of the current conjunction specifically, areas 6MA (Ruan et al., 2019) and 6D3 (Sigl et al., 2019) in the pre-SMA were significantly active (see Table 11). Figure 21 shows activation in the premotor cortex and a parameter estimate graph for area 6MA. No region (or set of regions) that could readily be identified as

the site of a phonological store was observed to be consistently active across both task phases. Critically, the activation along the left intraparietal sulcus (BA 7) in the results from this contrast —which some may interpreted as the location of the phonological store as it lies within the parietal lobe—was not observed in the results of the same contrast with auditory sequences in Experiment 1 (Chapter 3, Section 3.3.2.3). Instead, activation of various subdivisions across the occipital cortex further confirmed predictions that activity across both task phases could be explained in terms of motor planning and perceptual processing. Areas hOC3v and hOC4v (Rottschy et al., 2007) (h, human; OC, occipital cortex; v, ventral) are two distinct cytoarchitectonic areas located in the ventral portion of the occipital cortex. Topographically, areas hOC3v and hOC4v are proposed to represent the anatomical substrates of areas V3v and V4v (Rottschy et al., 2007). Areas in the lingual (BA 18 and hOc3v, V3v), occipito-fusiform (hOc4v, V4v) and superior occipital (BA 19) gyri, as well as the calcarine sulcus (hOc6, V6), were also active during both presentation and the temporal delay phase. BA 18 likely corresponds to the functionally defined V2 (prestriate) (Amunts et al., 2000) which receives information from the primary visual cortex (V1) and feeds forward to extrastriate areas such as V3 and V4. Both V2 and V3 carry afferent signals to V6 regarding motion (Galletti et al., 2001).

Figure 21

Parameter estimate and sagittal slice of MNI152 template image from results of Presentation | Temporal delay {T} conjunction in Experiment 2. Activation shown along the pre-central gyrus and parameter estimate plot area in 6MA (MNI: -6, 9, 51).

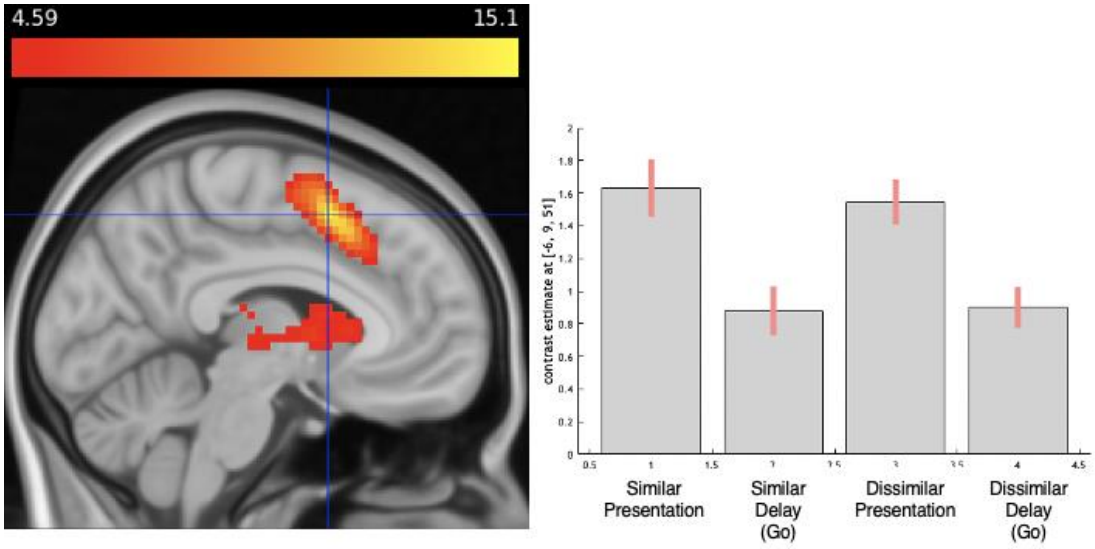


Figure 22

Comparison image across Experiments 1 and 2. Sagittal slices of MNI152 template image showing activation in the premotor cortex from results of the Presentation | Temporal delay {T} contrast.

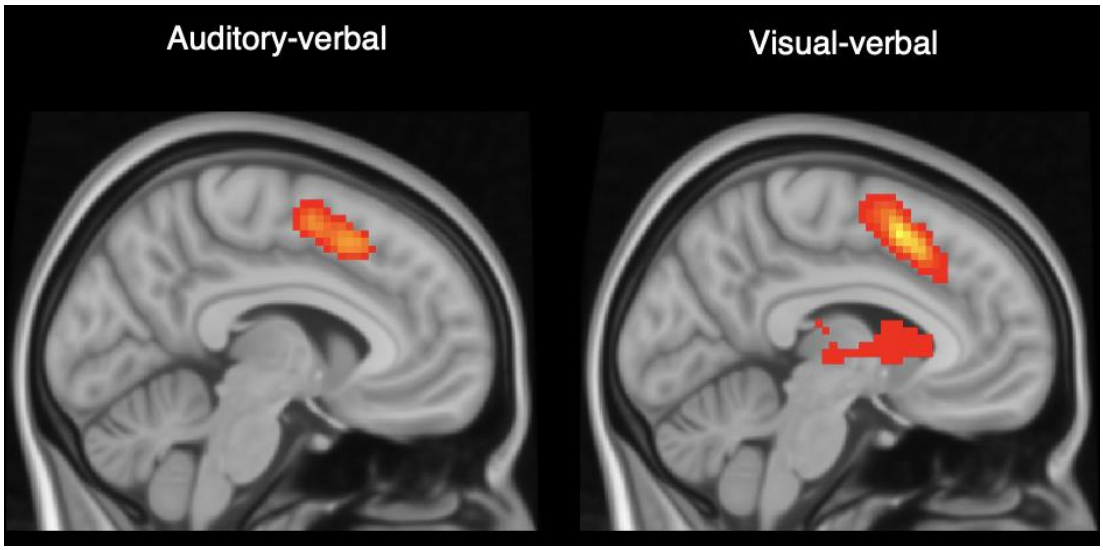


Table 11

Results of a factorial analysis of variance (ANOVA), Presentation | Temporal delay (Go) {T} conjunction in Experiment 2 (FWE-corrected for multiple comparisons, p<.05).

Gross Anatomy	F	Z	Coordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Frontal Lobe							
Prefrontal cortex							
Middle frontal gyrus	5.03	4.78	39	36	30	9	
Premotor Cortex							
Precentral gyrus	14.96	Inf	-6	9	51	6MA	58.1
Precentral gyrus	10.09	Inf	-39	3	33	6	
Precentral gyrus	7.85	7.05	24	6	54	6D3	74.7
Precentral gyrus	6.74	6.21	36	-9	60	6	
Precentral gyrus	6.63	6.11	36	-9	51	6	
Temporal lobe							
Basal Ganglia							
Caudate	6.42	5.94	-9	3	3		
Caudate	6.34	5.88	9	6	-3		
Head of Caudate	5.91	5.53	-6	18	6		
Thalamus	4.93	4.7	-3	-30	-3		
Thalamus	4.73	4.52	-6	-24	15		
Insular Cortex							
Insula	13.69	Inf	-27	24	0		
Insula	9.63	Inf	30	21	3	Id7	61.7
Parietal Lobe							
Intraparietal sulcus	4.75	4.54	-24	-63	42	7	
Occipital Lobe							
Lingual gyrus	10.93	Inf	-21	-84	-6	hOc3v [V3v]	50.3
Lingual gyrus	9.35	Inf	21	-87	-3	18	
Occipito-Fusiform gyrus	8.81	7.73	30	-78	-6	hOc4v [V4v]	71.3
Superior occipital gyrus	5.22	4.95	-21	-93	9	19	
Calcarine sulcus	5.14	4.88	-15	-72	12	hOc6 [V6]	68.9
Superior occipital gyrus	4.98	4.74	27	-90	12	19	
Cuneus	4.85	4.63	-24	-78	24	19	

4.3.2.5 Experiment 2: Repeating (Hebb) < Non-repeating (Filler).

A clear divergence was apparent between the trendlines for the Hebb and Filler conditions across cycles in the behavioural results (Section 4.3.1.2) but the noisiness of the data, particularly in the Hebb condition (which is based on fewer observations per cycle than the Filler condition) may have prevented detection of a reliable interaction. No significant differences in activation were observed as a function of repeating compared to non-repeating sequences. Extensive regions across the cerebral and cerebellar cortices were shown to be significantly active in results of the contrast between the temporal delays in Hebb and Filler conditions (Experiment 2B). However, the parameter estimate plots (Figures 23A - 23F) show that the results were driven by the difference between each trial-type (repeating or non-repeating) and their parametrically modulated regressors, similar to the results of Experiment 1B (cf. Chapter 3: Section 3.3.2.4).

Figure 23

Coronal slices of MNI152 template image showing activation across the cerebellum in lobules VIIA (Crus I & Crus II), VIIB, VIIIA and VIIIB in both the Filler (light blue) and Hebb (dark blue) conditions, in Experiment 2B. Parameter estimate plots showing activity evoked in Hebb Linear < Filler {T} contrast. A: Lobule HV, B: Lobule HVI, C: Lobule HVIIA (Crus I), D: HVIIA (Crus II), E: Lobule HVIIB, F: Lobule HVIIIA.

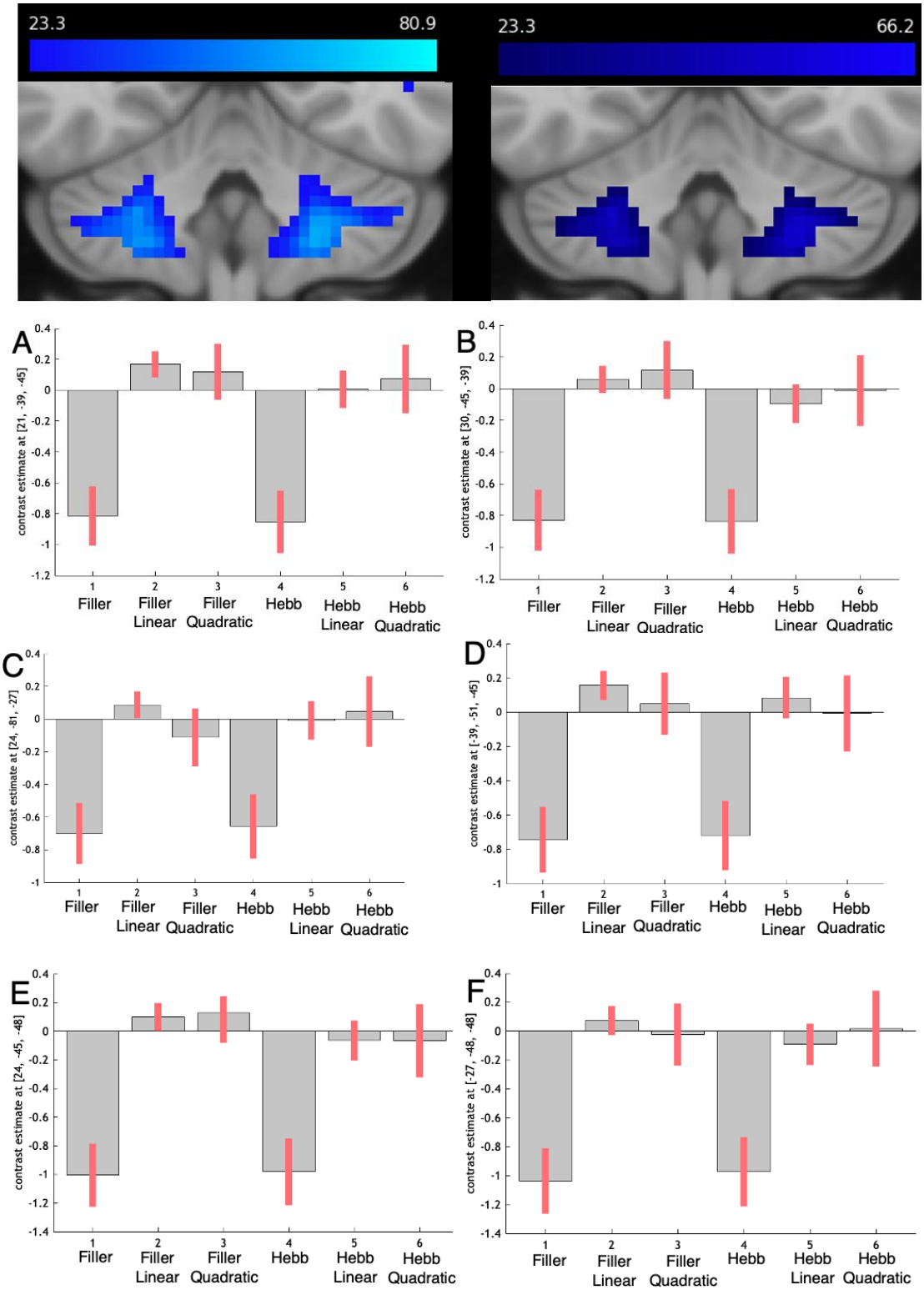


Table 12

*Results of a factorial analysis of variance (ANOVA), Repeating (Hebb) < Non-repeating (Filler) {T} contrast, (FWE-corrected for multiple comparisons, p<.05; *Prefrontal and Premotor cortices SVC; **Whole Cerebellum SVC; † Prefrontal, & Premotor cortices SVC and FWE-corrected for multiple comparisons, p< .05; ‡ Whole Cerebellum SVC and FWE-corrected for multiple comparisons, p <.05).*

Gross Anatomy	F	Z	Co-ordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Frontal Lobe							
Prefrontal							
Fronto-orbital cortex*	6.07	5.66	36	30	-15	11	
Fronto-orbital cortex*	6.05	5.64	33	27	-18	11	
Fronto-orbital cortex*	5.51	5.19	27	18	-21	11	
Anterior orbital gyrus	5.51	5.19	27	18	-21	11	
Lateral orbital gyrus*	5.25	4.97	-39	24	-18	11	
Superior frontal sulcus*	4.93	4.69	18	48	36	9	
Inferior frontal gyrus*	4.9	4.67	51	39	3	45	
Inferior frontal gyrus*	4.8	4.58	51	30	0	45	
Superior frontal sulcus*	4.61	4.41	12	24	60	8	
Lateral orbital gyrus*	4.52	4.33	-30	36	-15	11	
Superior frontal gyrus*	4.43	4.25	12	33	51	8	
Superior frontal gyrus*	4.35	4.18	-6	45	36	9	
Superior frontal gyrus*	4.32	4.15	12	42	48	9	
Superior frontal gyrus*	4.3	4.14	9	45	45	9	
Superior frontal sulcus *	4.22	4.07	-18	54	27	10	
Superior frontal gyrus*	4.19	4.04	-12	39	48	9	
Middle frontal gyrus*	4.07	3.93	-33	24	45	9	
Middle frontal gyrus*	4.05	3.91	-36	18	45	9	
Superior frontal gyrus*	3.96	3.83	21	54	27	9	
Superior frontal gyrus*	3.8	3.69	-18	45	36	9	
Middle frontal gyrus*	3.72	3.61	36	27	45	9	
Frontomarginal sulcus*	3.38	3.3	21	57	3	Fp1	72.1%
Posterior orbital gyrus*	3.34	3.26	-24	12	-24	11	
Inferior frontal gyrus,	3.27	3.2	-51	36	0	45	
Inferior precentral sulcus*	3.25	3.17	-60	0	6	44	
Superior frontal gyrus*	3.23	3.16	-9	24	57	8	
Inferior frontal gyrus,	3.22	3.15	48	24	12	44	
Premotor cortex							
Inferior Precentral gyrus	4.63	4.43	-45	-9	27	6	
Pre-central sulcus*	4.5	4.31	-48	-9	27	6	
Pre-central sulcus*	4.35	4.18	-42	-12	33	6	
Inferior precentral sulcus*	3.78	3.67	48	-9	30	6	
Temporal Lobe							
Insular cortex							
Insular cortex*	4.61	4.41	39	-9	-6		
Insular gyrus*	4.53	4.34	-42	-9	-3		

Gross Anatomy	F	Z	Coordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Insular cortex*	3.78	3.67	-33	-18	3		
Basal Ganglia							
Putamen	5.1	4.84	30	-9	9		
Caudate Nucleus	4.69	4.49	-21	-21	21		
Caudate Nucleus	4.65	4.45	27	-30	9		
Temporal Cortex							
Middle temporal gyrus	4.85	4.63	45	6	-42	38	
Middle temporal gyrus	4.74	4.53	57	-48	3	37	
Cingulate cortex							
Cingulate gyrus	6.24	5.79	-18	-51	33	31	
Cingulate gyrus	6.16	5.73	-21	-3	36	24	
Paracingulate gyrus*	5.95	5.56	-9	45	6	P32	51.7%
Paracingulate gyrus*	5.9	5.51	9	51	6	P32	75.1%
Paracingulate gyrus*	5.66	5.31	-3	54	21	P32	67.9%
Cingulate gyrus	4.98	4.74	30	-3	33	24	
Cingulate gyrus*	4.59	4.39	0	-15	36	24	
Cingulate gyrus*	4.34	4.18	0	33	15	24	
Parietal Lobe							
Superior parietal cortex	6.46	5.97	24	-42	-45	7	
Superior parietal cortex	5.48	5.17	3	-60	66	7	
Superior parietal cortex	5.47	5.16	-3	-51	69	5	
Inferior parietal cortex	5.35	5.05	-51	-57	30	Area PGa	58.6%
Postcentral gyrus	4.75	4.54	-21	-39	72	1,2,3	
Postcentral gyrus*	4.71	4.5	0	-33	72	4	
Superior parietal cortex*	4.55	4.36	-15	-33	39	5Ci	78.3%
Superior parietal cortex*	4.51	4.33	-9	-36	45	5	
Superior parietal cortex*	3.72	3.61	12	-33	45	5Ci	78.3%
Superior parietal cortex*	3.58	3.48	-3	-33	48	5	
Cerebellum							
Dorsal Dentate Nucleus **	7.33	6.65	-24	-69	-36		
Crus II**	6.02	5.62	-39	-51	-45		
Lobule V**	5.95	5.56	21	-39	-45		
Lobule VIIB**	5.92	5.53	24	-45	-48		
Lobule VIIB**	5.91	5.53	24	-51	-51		
Lobule VIIIB**	5.82	5.45	-21	-45	-48		
Lobule VIIIB**	5.8	5.44	-27	-48	-48		
Crus II**	5.39	5.09	-42	-66	-42		
Lobule VI**	5.3	5.01	30	-45	-39		
Crus I**	5.16	4.89	24	-81	-27		
Crus I**	4.91	4.68	-45	-66	-33		
Crus I**	4.78	4.57	27	-69	-36		
Crus I**	4.73	4.52	21	-72	-33		
Lobule VI**	4.49	4.31	-30	-42	-36		
Lobule IX**	4.23	4.08	12	-51	-54		

Gross Anatomy	F	Z	Coordinates (x, y, z)			BA	Probabilistic Cytoarchitecture (if available)
Crus I**	3.91	3.79	45	-48	-42		
Crus I**	3.88	3.76	45	-57	-42		
Crus II**	3.86	3.74	42	-51	-45		
Lobule V**	3.65	3.55	-21	-42	-24		
Lobule VIIB**	3.63	3.53	6	-75	-45		
Lobule VIIB**	3.62	3.52	-6	-78	-45		
Crus II**	3.49	3.4	36	-54	-45		
Crus II**	3.35	3.27	42	-69	-42		
Lobule V**	3.23	3.15	18	-51	-27		

4.4 Discussion

The results of present experiment support the view that visual-verbal serial-short term memory performance can be explained by recourse to articulatory planning and visual-perceptual processing. Phonological-store theory would suggest that the *same* region (or set of regions) separate to motor or perceptual regions should be observed during both presentation and temporal-delay rehearsal of sequences and regardless of presentation modality. Paramount to the perceptual-motor account however, no activation consistent with the notion of a phonological store was observed in either the visual or auditory studies. More generally, the common role of articulatory planning demonstrated across both experiments further confirms predictions of the perceptual-motor account.

To summarise the most important results from the present experiment, during the delay period (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]) activation was observed in the prefrontal cortex (area 9/46), precentral gyrus (BA 6), IFG (pars opercularis, BA 44) and cerebellum (lobules HVI and HVIIB). The common activation of regions[5] involved in goal-directed action planning during this task-phase with both

[5] Activation of the IFG (pOp, BA 44) was only observed at a corrected threshold across the temporal delay in the visual study.

visual (present experiment) and auditory sequences (Experiment 1) is in line with the view that articulatory rehearsal is engaged to support serial recall regardless of input modality (e.g., Jones et al., 2004). Although activation of BA 46 was not observed following FWE correction or SVC using a prefrontal cortex and premotor cortex mask in the results of the *Temporal delay (Go) > Temporal delay (No-Go)* contrast in the current experiment, an activation cluster (MNI: 42, 30, 27) in the middle frontal gyrus to BA 9, falls within area 9/46 (Pandya & Petrides, 1999) and is in close proximity to those localised by Hayter et al. (2007; MNI: 46, 32, 24). Co-ordinates localised to area 9/46 (Petrides & Pandya, 1999, 2002) in the auditory domain (MNI: -36, 33, 24 and 33, 45, 12) (Chapter 3, Section 3.3.2.1) were observed at a lower threshold ($p < .01$) in the results of the same contrast in the present experiment. It may be, then, that different areas of the prefrontal cortex are involved in the maintenance of high-level representations dependent on presentation modality.

In line with the hypotheses, the activation of BA 6 in the precentral gyrus, BA 44 (pars opercularis; pOp) in the IFG and area 9/46 in the middle frontal gyrus was observed during articulatory rehearsal of visually presented verbal sequences (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]). Activation along the superior frontal gyrus in BA 8 was observed during the temporal delay in this experiment (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]) but not in the results of the same contrast in the auditory experiment; its activation may, therefore, be specific to visual processing. In non-human primates, area 8 corresponds to the frontal eye fields (FEF) (Huerta et al., 1986, 1987) and although the specific location of the FEF in humans remains unclear, similar localisations have been suggested in humans in recent years (Percheron et al., 2015). Indeed, the co-ordinates (MNI: 33, 6, 60) in the results of this contrast are in close proximity to those observed by Mills (2017) (MNI: 24, 2, 58) in

a study on oculomotor behaviours. It is plausible that given the interconnections of the prefrontal cortex, the activation in the superior frontal gyrus (BA 8) reflects the operation of perceptual-motor mapping processes that may involve higher-order representations (Badre & D'Esposito 2007; Picard & Strick, 2001). For example, there is evidence that the region is part of the cascading hierarchy that departs posteriorly from the prefrontal cortex to the premotor cortex where predictive motor signals alter activity in perceptual regions and vice versa (Fuster, 1990, 2004; Hickok et al., 2011; Hickok, 2012; Miller & Cohen, 2001; Rauschecker, 2011; Rauschecker & Scott, 2009).

The results of the conjunction analysis (*Presentation* | *Temporal delay* [*Go*]) revealed that there were activations along the middle frontal gyrus (BA 9) and precentral gyrus (areas 6D3 and 6MA in the pre-SMA) during both the presentation and temporal delay phases of the task. These results are similar to those in the auditory experiment, indicating that the common mechanisms across modality during both presentation and retention of sequences lie within the prefrontal and premotor cortices. These results further indicate that the motor planning process begins during sequence presentation and continues during the delay period in the absence of any stimuli. Area 6MR (Amunts et al., 2019; Amunts & Zilles, 2015), another cytoarchitectonic subdivision of the pre-SMA, was also significantly active during presentation of visual-verbal sequences (*Presentation > Temporal delay [Go]).* This subdivision was also significantly active with auditory sequences but in that case only during the temporal delay (Chapter 3, Section 3.3.2.1). As the pre-SMA has been associated with higher-level abstract aspects of linguistic sequencing (Alario et al., 2006; Segaert et al., 2012) and the inhibition of overt vocalisations (Xue et al., 2008), comparison of pre-SMA activation across the two studies in the present thesis suggests that inhibition

may occur at different time points during verbal short-term memory tasks depending on modality.

Also active during the temporal delay in the present study was the IFG, in particular the pOp (BA 44) (*Temporal delay [Go] >Temporal delay [No-Go]*). This was also observed with auditory sequences in Experiment 1, though only at a lower threshold ($p < .01$; Chapter 3, Section 3.3.2.1). The activation of BA 44 (MNI: -54, 12, 6) is in close proximity to that observed in Papoutsi et al. (2009) (MNI: -54, 12, 12). These co-ordinates fall within the ventral portion of pOp which, according to Papoutsi et al. (2009), is involved in phonetic encoding. The activation of this region is also consistent with other research that has shown that the ventral region is involved in covert articulation, with greater activation in this region found in tasks involving phonetic compared to semantic retrieval (Heim et al., 2009). More specifically, the IFG has been shown to enable sub-lexical processing such as phoneme monitoring and sequencing which are integral sub-processes within covert articulatory rehearsal (Burton et al. 2000; Demonet et al. 1996; Poldrack et al. 1999; Zatorre et al. 1992, 1996). For example, in Indefrey and Levelt's (2004) model, the left IFG is part of a network related to syllabification, prior to the retrieval or compilation of the articulatory codes for syllables in the premotor cortex (BA 6). Alternatively, computational models of speech production propose that the role of the pOp, together with other regions within ventral premotor cortex, is to hold "speech sound maps", that is, representations of sounds and the motor programs by which they are generated (Guenther et al. 2006; Hickok & Poeppel, 2000, 2004, 2007). In this view, pOp acts as an integrative sensorimotor interface where the generation or retrieval of articulatory codes occurs.

An alternative interpretation of the activation of BA 44 could be the involvement of the region in an error-related feedback system for speech production. Rauschecker and Scott (2009) suggest that once information has been translated into motor-articulatory representations in BA 44, these are then communicated to the parietal lobe as an efference copy. Results of this contrast did show significant activation along the IPS and may indeed be reflective of an efference copy of motor representations or could reflect the connectivity of the anterior IPS to prefrontal regions and posterior IPS to early and higher visual regions (Bray et al., 2013; Mars et al., 2011; Uddin et al., 2010). Moreover, it has been suggested that the pOp (BA 44) is not involved solely in processing specific to speech but may also mediate higher-order movement planning (Binkofski et al., 2000). The proposal that pOp may be a domain-general region stems from early observations in non-human primates where F5, the homologue to human BA 44/45, was found to be active during observation and execution of a sequence of motor movements, demonstrating the relation between observation and action (Rizzolatti & Arbib, 1998).

The view of general-purpose motor planning regions supporting verbal serial recall is further substantiated by cerebellar activation in the results of the *Temporal delay (Go) > Temporal delay (No-Go)* contrast, where cerebellar lobules HVI and HVIIB showed similar patterns of activity to motor planning regions in the frontal cortex (see parameter estimate graphs in Figures 15A-C, 16A & 16C). These results are in line with previous research suggesting that cortical-cerebellar loops operate between cerebellar lobule HVI and motor/premotor cortices (Kelly & Strick, 2003) and the somatosensory cortex (O'Reilly et al., 2010) as well as between lobules HVI and HVIIB and the prefrontal cortex (Kreinen & Buckner, 2009). Such loops support the automatisation of motor skills and the results of this contrast are therefore in line

with the conceptualisation of verbal serial recall as a motor skill (e.g., Jones et al., 2004; Macken et al., 2015). Different clusters of cerebellar activation, bilaterally in lobule HVIIB, were observed during presentation of visual-verbal sequences (*Presentation > Temporal delay [Go])*. The co-activation of lobule HVIIB and the prefrontal cortex is again consistent with the notion of cortico-cerebellar loops (Kreinen & Buckner, 2009). Also, more active during presentation than the temporal-delay was Crus II of Lobule VIIA. Although Crus II is connected to the prefrontal cortex, no prefrontal regions were found to be active at a corrected threshold. The Crus II activation likely reflects the establishment of high-level representations or rules necessary for action selection and planning. If one makes the reasonable assumption that action selection, or the rules that guide action planning (would be established during presentation of the sequence as an articulatory trajectory is formed, this may explain why Crus II was active in the *Presentation > Temporal delay [Go]* contrast and exhibited a decrease in activation during the temporal delay (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]) contrast; cf. Figure 16B). During the presentation of auditory lists (Experiment 1), however, activation of cerebellar lobules HVIIB and Crus II during presentation were only observed at a lower threshold ($p < .01$). Similar to the interpretation of BA 44 activation at a lower threshold during the temporal delay in the auditory study, the activation of HVIIB and Crus II during auditory presentation at a lower threshold is not considered to be due to chance.

The discussion up to this point has focussed on the contribution of articulatory planning regions to visual-verbal serial short-term memory performance. Another key consideration, however, was to assess whether regions active during sequence presentation would demonstrate the operation of perceptual, or perceptual-motor mapping processes. As results of the *Presentation > Temporal delay (Go)* confirmed

predictions of the perceptual-motor account: That regions involved in auditory-perceptual processing and motor planning would be observed during sequence presentation that were not consistently active across both the presentation and temporal delay phases. It was predicted, then, that different regions involved in perceptual-processing would be observed during visual-verbal sequence presentation in comparison to auditory, but that similar regions of the premotor cortex may be revealed.

Consistent with the modality-specific predictions, activation in the lateral occipital cortex (LOC; hOc4lp specifically) was observed, in accordance with previous research demonstrating its involvement in the processing of object shapes (Grill-Spector, 2003; Larsson & Heeger, 2006). Region hOc4lp is also associated with tasks that require spatial location discrimination, visual attention, and visual tracking (Malikovic et al., 2016). Activation in this cytoarchitectonic sub-division therefore suggests that these aspects of visual processing occurred during the processing of visual-verbal sequences. Region hOc4lp was not observed in the results of any other contrasts in the current experiment and may indicate that the region is involved in visually-guided motor acts (Juypers et al., 1965; Pandya & Kupers, 1968) such as the timely assembly of visual stimuli into an articulatory plan (e.g., Macken et al., 2016).. Moreover, the co-activation of hOc4lp with the premotor, temporal, intraparietal and cerebellar (Crus II) cortices during presentation suggests that lateral occipital regions are potentially involved in a network related to the spatial and action-related processing that enable reading and language-related functions (Malikovic et al., 2016). The activation of an area in left middle temporal gyrus (BA 22) is consistent with research suggesting that this region is involved in the visual perception of words and letters (Davis & Gaskell, 2009; Fiez et al. 1999; Gow, 2012; Hickok & Poeppel 2004;

Indefrey & Levelt 2000). However, a letter-specific area in a region of the left extrastriate cortex (BA 37) (Flowers et al., 2004), lateral to the visual word-form area (VWFA), was not found to be active in the current experiment, contrary to previous claims that the region is active during the perception of letters. There is debate as to whether responses in occipito-temporal regions become selective to learned orthographic representations (Dehaene & Cohen, 2007, 2011) as, more generally, the left occipito-temporal cortex has been shown to be involved in the perceptual processing of visual features that are present in visual stimuli to differing degrees such as words, objects, letters and faces (Barton et al., 2010; Mei et al., 2010; Reinke et al., 2008).

Further evidence indicative of early visual processing was demonstrated through activation of the posterior fusiform gyrus (BA 19) during presentation of the visual-verbal sequences. Research has shown that such activation is attributable to reading-related activity and orthographic processing (Indefrey et al., 1997; Pugh et al., 1996). The activation in the posterior fusiform gyrus in the present experiment is also consistent with notions that the region contains the VWFA (FG2; Caspers et al., 2013). More generally, orthographic processing in the left ventral occipito-temporal cortex (in which the posterior fusiform gyrus falls) has been shown to be necessary for mapping visual forms to articulatory representations (Yarkoni et al., 2008) where progressive specificity for visual text-perception has been shown (Ben-Shachar et al., 2011; Brem, 2010; Maurer et al., 2006). Activation in this region is therefore consistent with predictions from the perceptual-motor account that modality-specific regions—different from those observed in the auditory domain—are involved in perceptual-motor mapping during verbal serial short-term memory. Moreover, the peak co-ordinate clusters in the premotor cortex observed during visual-verbal sequence

presentation were not observed in results of the visual conjunction analysis (Section 4.3.2.4) nor in results from any other contrast across modalities. This was also the case for peak co-ordinate clusters observed during auditory-verbal sequence presentation. These results further extend upon the perceptual-motor predictions and suggests that regions of the premotor cortex are attuned for modality specific perceptual-motor mapping.

In contrast to the results of Experiment 1, subdivisions of the cuneus (BA 17; BA 19 and V3) were observed to be more active during temporal-delay rehearsal of visual sequences as compared with presentation (*Temporal delay [Go] > Presentation*). The same contrast in the auditory domain revealed no significant activations. This finding may be related to the fact that such early visual areas have been shown to be active during visual imagery in the absence of any visual stimulation (Klein et al., 2000; Stokes et al., 2009, 2011). Instances of mental imagery can be understood as a particular instance of perceptual access (Kosslyn et al., 2001). For example, Stokes et al. (2009, 2011) observed activity across the LOC regardless of whether participants viewed the letters "X" or "O" or only imagined them. Such results suggest that visual imagery may evoke activation in subregions of the LOC different from those active during sensory-based perception. It may be, then, that subjects engage in visual imagery simultaneously with, or as part of, the articulatory planning of material that had been presented visually. Consistent with this interpretation, no significant activations of the LOC were observed in the same contrast in the auditory experiment.

Overall, the observed activations during the presentation of visual-verbal sequences across the occipitotemporal cortex, simultaneous with regions in the premotor cortex that were not observed in results of any other contrast suggest that a

form of perceptual-motor mapping is necessary in verbal serial short-term memory tasks. Results from this contrast across both experiments have demonstrated that areas within the premotor cortex were active only during sequence presentation. Moreover, peak co-ordinates of premotor cortex across modality differed to one another suggesting that premotor responses are dependent on stimulus modality. When these results are taken together with those of the *Presentation* | *Temporal delay* (*Go*) contrast across studies, results clearly indicate that verbal serial short-term memory phenomena can be attributed to the operation of perceptual processing and motor planning.

On the phonological-store theory, any region(s) to be identified as the phonological store would be expected to be active during presentation regardless of input-modality as well as over a delay period as the articulatory rehearsal process refreshes decaying items within the store, again regardless of modality. Based on the results of the current experiment, one might be tempted to suggest that the activation of the posterior IPS during both the presentation and delay phases (*Presentation* | *Temporal delay* [*Go*]) could be indicative of the operation of a phonological store. However, critically, the same activation was not observed in results of the same conjunction analysis with auditory sequences (Chapter 3, Section 3.3.2.3). Whilst the activation of this region may not be readily explained in terms of motor planning or perceptual processing either, the fact that it was not active consistently across both experiments violates one of the defining assumptions of the phonological store concept, namely, that the store holds phonological (and hence modality-independent) representations.

Original claims located the store in the left hemisphere, therefore, the activation observed along the right intraparietal sulcus (BA 7) that was greater during presentation in comparison to the temporal delay (*Presentation* > *Temporal delay*

[*Go*]) in the present experiment is inconsistent with previous attempts to localise the phonological store. Moreover, bilateral activation clusters observed in the intraparietal sulci (BA 7) during the temporal delay (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]) differed to those observed in the results of the same contrast in the auditory experiment (Chapter 3, Section 3.3.2.1). Whilst the possible involvement of the IPS in verbal serial short-term memory may warrant further investigation, the observation that different subdivisions were active at different task phases suggests strongly that none of the subregions are indicative of phonological store functioning or articulatory planning. As such, the activation of these regions is rather moot in the context of the main aims of the present thesis.

An investigation into the neural underpinnings of verbal sequence learning in the visual domain (Experiment 2B) revealed results similar to those presented in Chapter III with auditory stimuli (Experiment 1B). Changes in the BOLD-signal again appeared to reflect general task-set, rather than sequence-specific, learning. Although a main effect of sequence-type was observed in the behavioural analysis, indicating that sequence learning occurred, no corresponding excitability decreases were observed in the cerebellum. The common patterns shown in the parameter estimate graphs for cerebellar lobules across both Experiment 2B (visual) and Experiment 1B (auditory) suggest that the way in which the contrast was set up was unsuitable for examining the sequence learning effect or the hypothesised direction of activity was incorrect. It was assumed that parametric methods would be suitable for assessing the graded changes predicted as a function of learning given that such methods have been used successfully to investigate changing activity during the acquisition of other motor skills (Ramnani et al., 2000; Sakai et al., 2002; Toni et al., 2001). Results of *Temporal delay (Go) > Temporal delay (No-Go)* also showed that the hypothesised direction of

activation in Crus I and Crus II was not as predicted. Such results may instead reflect multiple types of cerebellar plasticity (D'Angelo, 2014, Ito, 2006): the cerebellum's role in learning and the basis of learning via changes at cellular and synaptic levels occurs jointly through different types of plasticity (some bi-directional) during various phases of learning. It may then be difficult to pre-empt how information processing changes in the cerebellum (for further discussion of these ideas, see Chapter 5).

Another possible reason as to why the predicted decrease in BOLD-signal was not observed (specifically for repeating sequences) is that both long-term depression (LTD) and long-term potentiation (LTP) are proposed to be necessary for the cerebellum to successfully act as a learning device (Ito, 2006). It may be, then, that *increases* in BOLD-signal may have occurred, in line with research demonstrating increases in cerebellar activity related to motor-sequence learning (Jueptner et al., 1997; Ramnani & Passingham, 2001; Sakai et al., 2002b). Future investigations into long-term verbal sequence learning should consider modelling data over time in a way that can explore bi-directional changes. There are also suggestions that during motor learning, dependency fluctuates between cortico-cerebellar and cortico-basal ganglia loops (Doyon & Benali, 2005; Hikosaka et al, 2002). Studies have shown combinations of excitability changes: Increases in the cerebellum and thalamus but decreases in the striatum have been observed as articulation rates increase (Riecker et al., 2006). As excitability decreases in prefrontal and premotor projecting lobules were observed for repeating and non-repeating sequences across both the current experiments, the possibility remains that excitability increases specific to repeating sequences may occur somewhere else in the network of regions between the cerebellum and cortex (e.g., thalamus or striatum; Miyachi et al., 2002; Riecker et al., 2006).

4.5 Conclusions

The results of the present experiment support the view that short-term verbal serial recall performance is supported primarily by articulatory planning. Regions in the frontal cortex (area 9/46, BA 6 and BA 44) and cerebellar lobules HVI and HVIIB were significantly active during rehearsal of visual-verbal sequences, as was the case with auditory sequences. Activation in BA 44 however, was only observed during the temporal delay following visual presentation and was not observed at an FWE corrected threshold. All of these regions have previously been argued to contribute either directly to motor planning, or to be involved in processing that informs the motor planning process. Activation of regions across the occipito-temporal cortex and the premotor cortex during sequence presentation have been interpreted as related to a perceptual-motor mapping process. As predicted, no activations consistent with the construct of a phonological store were observed in results from any contrasts in Experiment 2.

Although a Hebb repetition effect was also observed in the behavioural results, the fMRI results did not confirm the hypothesis pertaining to changes in the cerebellum as a function of sequence learning. Instead, the excitability decreases likely reflected general learning of a task-set. Discussion as to how the experimental design, scanning parameters, statistical modelling and analysis might be altered in future fMRI studies of Hebb sequence learning is included in Chapter 5.

CHAPTER V: GENERAL DISCUSSION

5.1 Overview of Empirical Findings

The experiments reported in the present thesis have provided neuroscientific evidence for the view that verbal serial short-term memory performance can be explained by recourse to articulatory-motor planning and perceptual processing, with no need to postulate, and indeed no evidence found for, a specialised phonological store. A comparison of the auditory (Chapter 3) and visual (Chapter 4) experiments demonstrated that the rehearsal of verbal sequences activates the prefrontal cortex (area 9/46), premotor cortex (BA 6), and cerebellar lobules HVI and HVIIB regardless of presentation modality. The same cluster of BA 44 activation was observed at a higher threshold in the visual study but only at a lower threshold (p = .01) in the auditory domain, interpretations of which will be offered below. The neuroimaging results extend previous cognitive-behavioural findings within the perceptual-motor framework and contribute to the growing body of work demonstrating how verbal short-term memory and sequence learning phenomena are rooted in general-purpose motor planning and modality-specific perceptual processing rather than supported by a non-motoric, modality-independent, storage unit. At odds with the phonological store theory, no region in the inferior parietal lobe was found to be active during both stimulus-presentation and the temporal delay . Rather, distinct subdivisions of the intraparietal sulcus (IPS) were active during the temporal delay with auditory sequences and during presentation and the temporal delay with visual sequences. Furthermore, neural activation specific to auditory presentation was observed in the premotor cortex, primary auditory cortex and planum temporale, whilst activation in

the premotor cortex, temporo-occipital-fusiform cortex, Crus II and lobule HVIIB was observed during visual presentation.

Experiments 1B (auditory) and 2B (visual) examined whether long-term learning of verbal sequences was associated with a down-regulation of the BOLD-signal in the cerebellum consistent with motor skill automatisation. Behavioural evidence of sequence learning was observed across both input-modalities but BOLD-signal changes across the course of Experiment B appeared to reflect general task-set, rather than sequence-specific, learning.

5.2 Verbal Serial Short-term Memory

For nearly 50 years, cognitive-behavioural research on verbal serial short-term memory has been influenced heavily by the concept of a specialised short-term phonological store (Baddeley & Hitch, 1974). The cognitive neuroscientific research presented in this thesis, however, has demonstrated that no single region consistent with the cognitive-psychological characteristics of the phonological store was observed across the cortex during auditory-verbal or visuo-verbal serial short-term memory. In contrast, the common activation of the premotor cortex during both sequence presentation and a temporal delay prior to recall is consistent with the view that articulatory motor planning underpins verbal serial short-term recall. Modality-specific activation during sequence presentation (*Presentation* > *Temporal delay*)—concurrent with regions in the premotor cortex that were not observed in results of the conjunction analysis (*Presentation* | *Temporal delay*) —suggests that the encoding of to-be-remembered items, regardless of presentation modality, involves a perceptual-motor mapping process. These results further confirmed predictions of the perceptual-motor approach as these activations were observed in the absence of the activation of

a region that could be plausibly interpreted as the location of a phonological store, thus contradicting the notion of a store holding abstract, modality-independent, representations.

The results across both modalities support and extend the cognitive-behavioural research by providing evidence that similar regions of the prefrontal and premotor cortices are active during the rehearsal of auditory- and visual-verbal sequences (*Temporal delay [Go] > Temporal delay [No-Go]*). In particular, similar clusters of activation localised to BA 9 along the middle frontal gyrus during the retention of auditory (MNI: -42, 21, 30) and visual sequences (MNI: -42, 30, 27) fall within area 9/46 as outlined by Petrides and Pandya (1999, 2002) and are likely related to the encoding of abstract information, monitored and maintained over the temporal delay necessary for achieving goals (Funahashi 2001; Fuster 1997; Petrides 1994). Similar clusters of activation along the middle frontal gyrus (BA 9) were also observed in results of the *Presentation | Temporal delay [Go]* conjunction (in the auditory [MNI: 42, 39, 30] and visual [MNI: 39, 36, 30] experiments). According to the adaptive coding model (Duncan, 2001), the prefrontal cortex is deemed capable of integrating almost any kind of information due to the plastic nature of prefrontal neurons and the consistent activation of BA 9 across both the sequence presentation phase and temporal delay may suggest that BA 9 is also involved in the maintenance of higher-order representations for action selection and planning.

The cascading hierarchy of the frontal lobe indicates that direct projections from the prefrontal cortex to the premotor cortex (Lu et al., 1994) suggest that higher-order representations in area 9/46 may be translated into motor plans in the premotor cortex. The similarity of activation clusters along the pre-central gyrus in the left premotor cortex (BA 6) during the rehearsal of both auditory-verbal (MNI: [-45, 0, 48]

[-48, -3, 54] [-42, 0, 27]) and visual-verbal (MNI: [-48, -3, 45) and [-42, 3, 30]) sequences—concurrent with proximal clusters in cerebellar lobule HVI (auditory [MNI: -27, -60, -30], visual [MNI: -24, -60, -30]) and HVIIB (auditory [MNI: 27, -72, -51], [30, -66, -51], visual [MNI: 27, -69, -48])—indicate the operation of closed loop circuits (Kelly & Strick, 2003; Lu et al., 2007) (Figures 17A-17C, Chapter 4). The proximity of cortical and cerebellar activations indicative of motor- planning across both, the auditory and visual experiments substantiates—from the perceptual-motor perspective—that verbal serial short-term memory performance is rooted in general-purpose motor planning. Further similarities in peak co-ordinate clusters in the premotor cortex were observed in the results of the *Presentation* | *Temporal delay* [*Go*] conjunction. These similarities were observed in the right hemisphere premotor cortex along the pre-central gyrus (auditory [MNI: 30, -6, 51] and visual [36, -9, 51]) and, more specifically, in the pre-supplementary motor area (pre-SMA).

Based on cytoarchitectonic properties of the region, subdivision 6MR (Amunts et al., 2019), 6MA (Ruan et al., 2019) and 6D3 (Sigl et al., 2019) all fall within the pre-SMA. Activation across modalities during both the presentation and temporal delay phases in area 6D3 (Sigl et al., 2019) occurred in different hemispheres (auditory [MNI: -21, -3, 54] and visual [MNI: 24, 6, 54]), whilst area 6MR (Amunts et al., 2019) was active during different task phases dependent on modality: auditory temporal delay (MNI: -6, 3, 57) and visual presentation (MNI: -6, 6, 60). In contrast, area 6MA (Ruan et al., 2019) [MNI: 6, 9, 51] was only found to be active in in the results of the visual conjunction (*Presentation* | *Temporal delay* [*Go*]). The role of specific subdivisions within the pre-SMA during distinct verbal short-term memory task phases are currently unclear.

Overall however, the results are consistent with the involvement of the pre-SMA in high-level planning processes via connections to the prefrontal cortex (Lu et al., 1994; Luppino et al., 1993). Higher level action plans are often related to execution but the pre-SMA has been implicated in the inhibition of vocal and manual outputs (Xue et al., 2008) as well as overt response inhibition during go/no-go tasks (Mostofsky et al., 2003; Sharp et al., 2010). Such results may be explained by recourse to the presumably uncontroversial assumption that articulatory planning is not equivalent to overt articulation as the latter requires additional processes such as the movement of orofacial musculature, inhalations and exhalations and so on that are not necessarily employed during covert rehearsal (or at least not to the same extent) (Palmer et al., 2001). Thus, it does not seem unreasonable to suppose that the pre-SMA served to inhibit the overt vocal execution of the articulatory plan in the present experiments. Activation of the pre-SMA is also modulated depending on how actions are specified (internal or external). When required to generate a word (overtly or covertly) from a semantic category (internally specified action), activity in the pre-SMA increases when compared to activity during an externally specified event such as word-reading (Alario et al., 2006; Crosson et al., 2001). Increased activity of the area is also observed during pseudoword repetition as a function of the complexity of the stimuli (Bohland & Guenther, 2006) and during covert action naming in comparison to overt (Kemeny et al., 2006). However, in a broader context of bodily action, specifically motor imagery (imagined finger tapping), increased activity was observed in the pre-SMA relative to observing finger tapping, and was observed in a conjunction between action-observation, motor imagery and execution (synchronous imitation) (Macuga & Frey, 2012). Such evidence suggests that the pre-SMA's involvement in effortful motor processing is not only not specific to the verbal domain

but related to different stages of motor action. It is not entirely clear, however, why different cytoarchitectonic subdivisions of the pre-SMA were active during different task phases in the current experiments. However, as the pre-SMA receives input from the prefrontal cortex, activation of different cytoarchitectonic areas could be related to the transformation of abstract representations into motor commands during and following perceptual input. Whilst these differences should be acknowledged, the domain-generality of the pre-SMA should also be considered given its co-activation in a larger network spanning cortical and sub-cortical regions for motor planning. A pre-SMA-thalamic-cerebellar pathway is proposed to facilitate timing, prediction and sequencing of events that involves the caudate (Kotz & Schwartze, 2010) and is consistent with the basal ganglia being involved in the innervations of vocal tract muscles (Brendel et al., 2010). In the conjunction analysis for the auditory experiment, the thalamus and area 6D3 were active whereas during the visual conjunction analysis, the caudate, and areas 6MA and 6D3 were active. These results suggest that the operation of various regions across cortical, sub-cortical and cerebellar regions contribute to motor planning.

Based on the results of Experiment A across both the auditory and visual modalities, the cerebellum's involvement was consistent with the traditional view of its role in monitoring, controlling and executing motor actions (Babinski, 1902; Holmes, 1939; Manto et al., 2012) via temporal organisation and sequencing dynamics of co-ordinated action (Barlow, 2002; Braitenberg et al., 1997; Ito, 1993, 2005; Thach et al., 1992; Wolpert & Miall, 1996). Particularly relevant to the formation of an articulatory trajectory is how the cerebellum exerts sensorimotor control and co-ordinates vocal tract and laryngeal movements as well as respiration during overt speech production (Ackermann et al., 2008). Proximal clusters of activation in the

cerebellum across modalities highlights the centrality of articulatory planning during verbal serial short-term memory. In particular, lobules HVI and HVIIB were active during the temporal-delay retention of both auditory and visual sequences (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]). Bilateral activation in lobule HVI was observed during the retention of both auditory and visual sequences, but similar clusters were observed only in the left hemisphere (auditory [MNI: -27, -60, -30] and visual [MNI: -24, -60, -30]). The temporal delay in the present experiments did not require overt vocal rehearsal but activation observed in lobule HVI (MNI: 6, -78, -22) during temporal-delay rehearsal of auditory-verbal sequences (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]) was in close proximity to an activation cluster in lobule HVI (MNI: 6, -75, -24) in Hayter et al. (2007) during a conjunction of "Add" and "Repeat" conditions: Subjects in that study were required to add a presented digit to the immediately preceding digit in the former condition or repeat each number after hearing it in the latter but both tasks required overt vocal responses. Taken together, results from both the current experiments indicate that lobule HVI facilitates the planning and execution of verbal sequences. These results are in line with the view that lobule HVI generates internal models for the vocal tract (Callan et al., 2007) as it is proposed to receive afferent information from oral and facial musculature (Stoodley & Schmahmann, 2010). This information provides the cerebellum with sensory information regarding the state of speech effectors and is consistent with propositions that lobule HVI contains the lip and tongue areas of a sensorimotor homunculus (Buckner et al., 2011; Manni & Petrosini, 2004; Mottolese et al., 2013).

The role of lobule HVI for repetitive and consecutive actions is also observable during non-speech motor-related tasks. Durisko and Fiez (2010) demonstrated that activation of lobule HVI occurred not only during covert and overt speech in a verbal

short-term memory task but also during covert and overt tapping. These results are consistent with the traditional role of the cerebellum in co-ordinating actions via sequencing and timing and, when taken together with the results of Hayter et al. (2007) and activation of lobule HVI in the current experiments, suggest that the lobule could contribute to domain-general overt and covert action planning. Similarly, the activation of lobule HVIIB is also implicated in a closed loop circuit that supports motor planning. During the temporal delay (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]), similar clusters of activation were observed in lobule HVIIB during auditory retention (MNI: 27, -72, -51 and 30, -66, -51) and visual retention (MNI: 27, -69, -48). Activation was observed bilaterally during visual retention but was right lateralised during auditory retention. The results from the current experiments challenge the view that a temporo-parieto-cerebellar loop (involving lobule HVIIB) supports a phonological store (Desmond et al., 1997). Contrary to phonological-store theory, lobule HVIIB does not exhibit consistent activation across all phases of verbal short-term memory tasks: In the present experiments, lobule HVIIB activation was observed during both presentation and temporal delay phases with visual sequences but only during the temporal delay with auditory sequences. The role of lobule HVIIB in articulatory planning is further supported by results of another delayed serial recall task where activation of lobule HVIIB was observed across a temporal delay prior to recall (Durisko & Fiez, 2010). The fact that in no study has lobule HVIIB been found to be active across all phases contradicts the notion that it could support phonological storage (cf Chapter 1, Section 1.5, Table 1): Any correlate of the phonological store should be active throughout all phases of a verbal serial short-term memory task.

The perceptual-motor approach asserts that the order and timing of items embodied within a motor plan contribute to the creation of a motor object, binding

otherwise discrete items together in a forward order (Jones & Macken, 2018). The activation of cerebellar lobules shown to project to the premotor cortex may, therefore, reflect the role of the cerebellum in feedforward control related to timing. If the cerebellum acts as a timing device (D'Angelo & De Zeeuw, 2009; Knolle et al., 2013) it may be involved in planning articulations as online and feedforward commands rapidly aid the sequencing of syllables into larger and coherent utterances (Ackermann, 2008). More generally, this view of the cerebellum is consistent with its supposed role in the planning of other behaviours executed via different bodily effectors that also follow sequential structures (*if this, then that* etc).

5.2.1 Unexpected results

Given the discussion of motor objects and representational units, it was particularly surprising that some predictions in the hypotheses were not met. It was predicted that activation of the inferior frontal gyrus (pars opercularis; pOp) would be observed in the results of the *Temporal delay* (*Go*) > *Temporal delay* (*No-Go*) contrast across both modalities. This assumption was partly based on the consistent co-activation of the area with the premotor cortex (BA 6) in the verbal short-term memory studies reviewed in Chapter 1, all of which used visual presentation (cf. Section 1.5, Table 1; Awh et al., 1996; Chein & Fiez, 2001; Fiez et al., 1996 Paulesu et al., 1993; Salmon et al., 1996). A further role of pOp in the generation of articulatory plans during both covert and overt rehearsal (Callan et al., 2006; Indefrey and Levelt, 2004; Papathanassiou et al., 2000) suggests that the region is involved in sub-lexical processing where segmentation is required (Burton et al., 2000; Zatorre et al., 1996). For this reason, activation of pOp was predicted across the temporal delays for both input modalities, contrary to the data.

Instead, POp was only observed across the visual delay period at a corrected threshold ([MNI: -51, 9, 18], cytoarchitectonic probability 54.2% and MNI: -54, 12, 6, cytoarchitectonic probability 66.5%). The co-activation of pOp with BA 8 and BA 9 along the superior frontal gyrus and BA 6 in the pre-central gyrus is consistent with the existence of structural connections between the regions (Ford et al., 2010). A cluster of activations originating from the former set of co-ordinates was observed at a lower threshold ($p < .01$) across the delay period during the rehearsal of auditory-verbal sequences. No activation was observed to originate from the latter set of co-ordinates, however, and could suggest that the sub-regions within pOp respond differently to motor plans requiring sensorimotor integration in different modalities. Given that the planum temporale is proposed to support auditory-speech integration (Hickok et al., 2009) and was only observed in the auditory experiment, it is possible that increased activation of pOp during temporal-delay retention of visual-verbal sequences reflects increased involvement of pOp following sensorimotor integration during presentation of visual-verbal sequences. Moreover, the significant activations of pOp during the temporal-delay retention of visual-verbal sequences were observed following FWE-correction or SVC of the frontal cortex. The use of an SVC specific to the IFG may have revealed significant activation clusters in the auditory experiment.

A second result, related to the direction of activity in Crus I and Crus II across the temporal delay (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]) did not meet predictions but instead indicated that learning of higher-order representations related to motor processing occurred on a shorter timescale than originally expected. The primary hypotheses were based on the assumption that forms of motor sequence learning involve learning rules that dictate movement (Tanji & Hoshi, 2001). These rules (higher-order representations) originate in prefrontal cortex and are translated to

motor commands in the premotor cortex as part of a goal-directed action schema. As Crus I and Crus II receive direct projections from area 9/46, similar patterns of activations—positive increases above baseline observable in parameter estimate graphs—across these regions were predicted concurrent with those regions facilitating motor-planning during the temporal delay in the short-term experiment. Whilst increased activation of area 9/46 relative to baseline was consistent with predictions, a decrease of BOLD-signal in Crus I and Crus II appeared to occur rapidly in the short-term serial recall task.

Indeed, activation of area 9/46, Crus I and Crus II in the results from multiple studies substantiate the view that a cerebellar-prefrontal cortex loop supports the acquisition of cognitive skills (higher-order representations) and subsequent action. For instance, Hayter et al. (2007) found that prefrontal cortex area 9/46 and cerebellar Crus II were significantly active during the "Add" condition of their study (Chapter 3, Section 3.4). A similarity was noted in location of Crus II activation (MNI: 6, -84, -34; Hayter et al., 2007) during their Add condition and the rehearsal of visual-verbal stimuli (MNI: 9 -78, 33) in the current Experiment 2A (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*]). A similar retention process was necessitated in both tasks but the task material differed (numerical vs. verbal) and the "Add" condition required subjects to continuously update computations to subsequently produce an overt verbal response whilst the present experiments required covert maintenance of a verbal sequence. Although proximal areas within Crus II were found to be active, it may be inferred that the differing patterns of activity likely reflect the sensitivity of Crus II in its forward-modelling capacity for higher-order representations. This is supported by evidence suggesting that the prefrontal cortex is able to integrate almost any kind of information (Duncan, 2001).

Other research has also suggested that areas of Crus II respond differently to sequence familiarity based on the acquisition of forward models. In comparing activation during the execution of oculomotor sequences without any instruction, Mills et al. (*in prep*) observed activation in Crus I and Crus II showing that novel sequences in comparison to familiar ones evoked greater BOLD-signal activity. Furthermore, graded changes in line with incremental increases in sequence familiarity (one repetition vs. three) showed that activity was greater for one repetition and declined as repetitions increased. A paravermal area of Crus II showed similar patterns of decreasing activity as Crus I, however a more lateral region in Crus II demonstrated greater decreases for novel sequences than familiar, which contradicted physiological predictions. As their results showed such patterns of activity, Mills et al. (*in prep*) suggested that different areas of Crus II were modulated by sequence familiarity based on the acquisition of forward models and more globally that the patterns of cerebellar activity are not yet entirely clear. It may be that Crus I and Crus II maintain general roles in higher-order representations unrelated to specific bodily effectors and a combination of both long-term potentiation (LTP) and long-term depression (LTD) — the strengthening and decreasing strength between synapses —could underpin learning. This view is in keeping with previous physiological work that has demonstrated that changes in excitability within the striatum are responsible for skill learning (Lehéricy et al., 2005; Poldrack et al., 2005).

When comparing the differences of Crus II activation to that of Hayter et al. (2007), it is likely that the learning of the higher-level representations related to action planning and selection necessary for achieving goals (via error feedback in forward models in Crus II) may have occurred on a shorter timescale than predicted. The current predictions were that a decrease of the BOLD-signal in the cerebellum would

be observed across the course of the long-term sequence learning experiment (Experiment B). However, increased activation of Crus II above baseline was observed during the presentation of visual-verbal sequences, and when informally comparing the parameter estimates of Crus II during the presentation phase (*Presentation* > *Temporal delay* [*Go*], Chapter 4, Section 4.3.2.2, Figure 18B) and the temporal delay (*Temporal delay* [*Go*] > *Temporal delay* (*No-Go*], Chapter 4, Section 4.3.2.1, Figure 16B), it is apparent that the BOLD-signal in Crus II changed depending on task phase. This comparison indicates that learning of higher-level representations necessary for the rehearsal of visual-verbal sequences (action selection and planning) occurred rapidly, prior to the temporal delay. To assess whether physiological changes related to learning of such representations did occur during the presentation phase, future investigations could model data in a similar way to the temporal delay period. The assumption that learning of higher-level representations occurred early on is further supported by the fact that Crus II activation was absent from the *Presentation* | *Temporal delay* (*Go*) conjunction analyses across both modalities. Neither were the co-ordinates of Crus II activation during visual presentation or Crus I and Crus II co-ordinates during auditory retention active at a lower threshold during auditory presentation. In fact, no Crus I or Crus II activation was observed during auditory presentation including the co-ordinates of Crus II that were active during visual retention, even at a lower threshold. Overall, these results indicate that different areas within Crus II respond to different task phases during verbal short-term memory tasks.

In contrast to the current results, evidence has shown that positive increases in Crus I and Crus II activation were also observed for the processing of both first-order (Balsters and Ramnani, 2011) and second-order rules (Balsters et al., 2013) where activity time-locked specifically to processing of rules (symbolic instruction cues) for

subsequent action (movement of individual fingers in a button press task) was measured. Activity time-locked to processing of first-order rules (symbolic instruction cues) showed activation in prefrontal area 9/46 and Crus I (Balsters & Ramnani, 2011) whilst activity for second-order rules extended from Crus I to Crus II (Balsters et al., 2013). As a result of receiving feedback regarding the accuracy of responses, activity in Crus I decreased (Balsters & Ramnani, 2011). Taken together, not only do the results indicate the ability to acquire associations between rules and action, they also suggest that the cerebellum is a candidate area for the automation of cognitive operations which guide motor execution through acquisition and storage of forward models (Balsters & Ramnani, 2008; 2011; Balsters et al., 2013).

5.2.2 The search for a non-existent phonological store?

The goal of the present research was not only to assess the extent to which motor planning and perceptual processing underpins short-term verbal serial recall but also to examine whether evidence of a region consistent with the characteristics of a phonological store would be found. The store is proposed to exist outside any speech or language architecture in the brain and must be activated when auditory-verbal information is being processed as such information enjoys obligatory access to the store. The review of regions in Chapter 1, Section 1.5, outlined why previously proposed regions were disregarded as the possible location of the store. Importantly, no region across the auditory and visual experiments was observed that could, according to the theory, plausibly be identified as a neural substrate of the phonological store.

Instead, the results revealed activation of modality-specific perceptual processing regions simultaneous with regions of the premotor cortex that were not observed specific to the presentation phase of the verbal short-term serial recall tasks.

Results from the *Presentation > Temporal delay (Go)* contrast in the auditory experiment demonstrated distinct regions of activation bilaterally in the auditory cortex (BA 41), along the superior temporal sulci (STS) posteriorly in the right (MNI: 66, -30, 3) and left hemispheres (MNI: -57, -21, 0). The latter co-ordinates are likely to lie within the planum temporale (PT) and are consistent with suggestions that the region is commonly activated by auditory inputs and covert articulation (Hickok et al., 2003, 2009). As noted in Chapter 3, the activation cluster in the PT observed during auditory sequence presentation (MNI: -57, -21, 0) was close in location to that observed by McGettigan et al. (2011) during presentation of auditory stimuli (MNI: -54, -18, 6) and rehearsal (MNI: -51, -28, 3) of their high vs. low syllable load comparison in a delayed pseudoword repetition task. The fact that activation of the region in the present auditory experiment was only observed during presentation is consistent with suggestions that the region functions as an auditory-motor interface; indeed, activation of the same region was not observed during visual-verbal sequence presentation. Instead, activation in the right lateral occipital cortex (hOc4lp; Malikovic et al., 2016), bilateral fusiform gyrus (BA 19), left posterior middle temporal sulcus (MNI: -48, -48, 6) and right STS (MNI: 51, -42, 9) were observed with visual-verbal sequences. Activation of hOc4lp during visual sequence presentation is consistent with the region's involvement in spatial location discrimination, visual attention and visual tracking (Malikovic et al., 2016). Furthermore, activation of the posterior fusiform gyrus (BA 19) is consistent with reading-related activity and orthographic processing (Indefrey et al., 1997; Pugh et al., 1996) in the region and the mapping of visual forms to articulatory representations (Yarkoni et al., 2008). Overall, it can be suggested that activations observed across the temporal and occipital cortices during the presentation of auditory-verbal and visual-verbal sequences are in agreement with the view that an

appeal to motor planning and modality-specific perceptual processing is sufficient to account for verbal serial short-term memory performance.

Assessing activation specific to task phase (presentation, temporal delay, recall) was in part inspired by the study of Chein and Fiez (2001) where activation of area 7/40 in the parietal cortex was observed during presentation only. Critically, activation of the phonological store should be observable during each task phase across both modalities but the results of the set of contrasts conducted here (*Temporal delay* [*Go*] > *Temporal delay* [*No-Go*], *Presentation* > *Temporal delay* [*Go*] and *Presentation* | *Temporal delay* [*Go*]) confirmed the perceptual-motor account's view that no such store exists. Instead the activations observed across the inferior parietal lobe can more readily be attributed to perceptual processing and motor planning. Moreover, activation of BA 40 in the inferior parietal lobe was not observed across any results from either the auditory or visual experiments. Instead, the majority of the inferior parietal lobe (IPL) results centre in and around the intraparietal sulcus (IPS). The IPS in macaque monkeys contains areas proposed to support sensorimotor integration (Andersen & Buneo, 2002; Colby & Goldberg, 1999), consistent with connections to the inferior parietal lobe (IPL) from the ventral premotor cortex (Petrides & Pandya, 1984). Specific to speech planning and production, and analogous to the articulatory rehearsal process, somatosensory state and error maps are supported by the IPL and are assumed to provide feedback for such processes (Bohland & Guenther 2006; Guenther, 2006). During the temporal-delay retention of both auditory and visual sequences (*Temporal delay [Go] > Temporal delay [No-Go]*) clusters of activation along the left IPS (auditory: MNI: -27, -57, 42, and visual: MNI: -24, -63, 42) and could be related to the efference copies of motor plans included in models of speech production (Rauschecker & Scott, 2009). These results are consistent with

others showing that the right dorsal premotor cortex is functionally connected to left IPS during the maintenance of a verbal sequence (Cairo et al., 2004).

All other co-ordinate clusters during the temporal delay differed according to presentation modality. Specific cytoarchitectonic results were observed in the auditory domain in the anterior part of the medial wall along the IPS, hIP3 (50.1%, MNI: 33, -63, 51) and the lateral wall of the IPS, hIP6 (51.4%, MNI: 30, -60, 39) (Scheperjans et al., 2005; 2007). Area hIP6 has been shown to be involved in action inhibition (Richter et al., 2019) and it is likely therefore that in the context of the verbal short-term memory experiments, hIP6 activation reflected the inhibition of overt vocalisations during the maintenance of auditory-verbal sequences. The exact same region was not observed during temporal-delay retention of visual sequences but one of the three co-ordinate clusters (MNI: 30, -60, 30) observed during visual presentation (*Presentation* > *Temporal delay* [*Go*]) was proximal to that of hIP6 in the results of the auditory *Temporal delay* [*Go*] > *Temporal delay* [*No-Go*] contrast (MNI: 33, -63, 51) and may fulfil a similar function. Results specific to the presentation of auditory and visual sequences (*Presentation* > *Temporal delay* [*Go*]) showed that activation was only observed along the IPS during presentation of visual items and was in the right hemisphere. On a phonological store-based account, the activation would not be considered indicative of the store as the store was proposed to be located in the left hemisphere. Another co-ordinate cluster observed during visual presentation (MNI: 30, -51, 48) was close to MNI: 30, -54, 48 during the retention of visual-verbal sequences. Critically, neither these co-ordinates, nor similar ones, were observed across the results from the auditory experiment, indicating that it could not be interpreted as the location of a phonological store. In the conjunction analysis (*Presentation* | *Temporal delay* [*Go*]), again the only active cluster was observed in

the visual experiment and was along the intraparietal sulcus (MNI: -24, -63, 42). This cluster was close to one observed to be active during auditory retention (MNI: -27, -57, 42).

Overall, the results from the contrasts demonstrate that whilst there were some proximal similarities in clusters of activation, no single region was observed to be consistently active during both presentation and retention of both auditory- and visual-verbal sequences. As discussed in Chapter 4, the involvement of the IPS in verbal serial short-term memory warrants further investigation. To further assess the role of the intraparietal sulcus in the different phases of a verbal serial short-term memory task, small volume corrections (SVC) could be applied at the analysis stage if keeping the experimental design the same. I turn now to discuss how the present results fit within the broader context of research reconceptualising the understanding of verbal serial short-term memory.

5.2.3 Representations within the perceptual-motor framework

Whilst the discussion up to this point has focussed on general purpose functions of motor planning facilitated by regions across the cerebral and cerebellar cortices (premotor cortex, pre-SMA, cerebellar lobules HVI and HVIIB), it is necessary to consider whether behavioural effects such as the 'phonological' similarity effect are also observed outside the verbal domain and can be explained with recourse to motor representations. Converging lines of evidence indicate that effects of 'phonological' similarity in the verbal domain are—at least when articulatory planning can be engaged— the result of articulatory, not phonological, errors (Acheson & Macdonald, 2009; Ellis, 1980; Jones et al., 2004; MacKay, 1970; Nooteboom, 1967; Page et al., 2007). Other evidence further supports the assumption that the root of the 'phonological' similarity effect lies within errors of motor planning and execution as

an equivalent effect is observed with manual hand gestures in deaf signers. Signs that share either handshape, location in space, movement, or palm orientation are considered similar within sign language. The serial recall of signs is lower for signs that are similar in comparison to dissimilar. This similarity effect, however, like the 'phonological' similarity effect in verbal serial recall, disappears when motor planning is impeded: When signers are required to repeat simple task-irrelevant movements of the hands, the sign-similarity effect is eliminated (Wilson & Emmorey, 1997). Thus, in both spoken and signed languages, under visual presentation, suppression of motor planning abolishes the 'phonological'/sign similarity effect (Baddeley 1986; Wilson & Emmorey, 1997). Performance is also modulated by the length of the to-be-remembered signed stimuli (Wilson & Emmorey, 1998). The sign-length effect is therefore akin to the word-length effect found with verbal stimuli in hearing subjects. Again, under manual suppression, the sign-length effect is abolished, therefore suggesting that the length effect originates from a 'sub-manual' rehearsal process. These results suggest that a sub-manual rehearsal process operates in a very similar way to the subvocal articulatory process used by hearing subjects with verbal stimuli (Wilson, 2001). Consistent with this, the left temporo-parietal junction, including the posterior superior temporal sulcus, supramarginal gyrus and planum temporale, have been found to be active during the perception of both signed and spoken languages (Hickok et al., 2003; Jacquemot et al., 2003; Wise et al., 2001).

An object-oriented action view of verbal serial short-term memory maintains that verbal information is apprehended and manipulated by the vocal effectors in the same general way as other kinds of input are apprehended and manipulated by other kinds of motor effectors (Jones & Macken, 2018). Goal-directed actions (movement, or the planning required for movement) are described as object-oriented as obligatory

perceptual processes organise environmental input (related to material and task) into perceptual objects that can be apprehended and manipulated by motor systems (Bregman 1990; Scholl, 2001). In this view, motor-action can therefore benefit from the obligatory access of the products of perceptual processes to motor systems (e.g., Rizzolatti & Luppino, 2001). For example, both action-planning related to a solid object and that related to a verbal 'object' should be considered in the same way as they are both perceptual-motor task settings and require the integration of perceptual and motor information. Where visual representations typically guide goal-directed action in regard to a solid object, both visual and auditory representations typically aid the transformation of verbal input for subsequent action given how representations enter the speech and language systems (written and spoken).

Of particular relevance to the issue of perceptual-motor mapping within the present thesis was the examination of the differences in neural activation according to presentation modality. Regions found to be active during sequence presentation in the auditory domain included the planum temporale whilst visuomotor integration was associated with regions across the occipitotemporal cortex. The formation of an auditory object is based on grouping over time (*sequential streaming*; Bregman, 1990), whereas visual object formation relies on spatial substrates (Scholl, 2001). The formation of the perceptual objects is thought to be an automatic, obligatory process but the process of mapping a perceptual object onto an appropriate motor plan can be affected by task requirements (Taylor et al., 2015). One consequence of this is that a rehearsal process best supports recall of early parts of a to-be-recalled sequence, as supported by studies in which participants are asked to rehearse overtly (Tan & Ward, 2002). Although serial recall is generally better for the start of a sequence, the studies that acted as the springboard for the perceptual-motor approach highlighted modality-

specific differences, where recall of auditory sequences benefits from the salience of the boundary of an auditory object (e.g., the primacy effect; Jones et al., 2006 and the recency effect; Jones et al., 2006, 2004; Nicholls & Jones, 2002). Other work within the framework has capitalised on the finding that when item presentation alternates from one voice or ear to another successively, sequences are more poorly recalled in comparison to when all items are presented in the same voice/to both ears (Greene, 1991; Hughes et al., 2009, 2011, 2016). It was argued that the alternating sequences were perceptually organised automatically into by-voice/ear objects thereby hindering the process of mapping the to-be-remembered items onto a motor plan that would faithfully replicate the order of items as-presented (Hughes et al., 2009, 2016).

Because the perceptual-motor approach asserts that individual items within a verbal sequence become bound into a single articulatory object, the view is at odds with the common notion that speech is represented in the brain in terms of phonemes. It has been argued that the idea that the brain represents speech in phonemes—discrete, static, context-independent, abstract entities—that map roughly onto the letters that make up written language is a trap that some psychologists have fallen into because of the invention of written languages (Macken et al., 2015). That is, psychologists have been seduced by the separability of letters in written language to the extent that they have assumed that there are corresponding units in the mind for processing the kind of material (speech) that the written form was invented to represent 'on paper'. This classical phonemic view of the nature of speech stems from literacy learning via segmental representations instead of the actual characteristics of speech sounds and the action processes by which they are generated (Port, 2007; Wray, 2014).

Critics of a phonemic view of speech processing point towards the syllable as a more relevant unit (Massaro, 1972; Savin & Bever, 1970). It has been suggested that

syllables may be coded as “motor chunks” (Hickok, 2013), also known as a mental syllabary (Cholin et al., 2006; Levelt et al., 1999; Levelt & Wheeldon, 1994). For example, Cholin et al. (2006) suggest that frequency effects in articulating particular syllables or sequences of syllables shapes experience which contributes to the generation of one’s mental syllabary. They suggest that more frequently encountered sequences will be more “strongly chunked” in the motor system, meaning that less segmental guidance is necessary. This then begs the question of how useful a phoneme-sized unit is. Hickok (2013) proposes that holistic chunks—a syllable or a sequence of syllables—does not require information of internal segmental structure and the order of phonemes is coded at a higher-level. What are conceptualised as ‘phonemes’ can actually be recovered based on higher-order (syllabic) processing if a task requires it, but phonemic level processing is not an essential, nor instrumental part of speech processing. Instead, a ‘phoneme’ is conceptualised as a somatosensory code than a solely auditory code. A somatosensory code embodies the physical-articulatory and subsequent acoustic properties regarding the state of the vocal tract and what are described as somatosensory consequences such as a relaxed or tort larynx, position of lips, teeth and pressure of tongue on the alveolar ridge. This view maintains some similarities to that of Liberman et al. (1967) and Browman and Goldstein (1995) who assert that the perception of speech is dependent on the mapping between acoustic and articulatory features. The representational unit is therefore not solely defined by sound nor by articulation but the state of the vocal tract and oral apparatus. If a phoneme has no acoustic realisation, then it cannot be coded in auditory space and acoustic properties alone cannot define the unit.

5.3 The Wider Context of the Perceptual-Motor Framework

The current fMRI experiments were the first to be conducted explicitly within the perceptual-motor framework (Hughes et al., 2009; Jones et al., 2004) though as noted throughout the thesis, the present approach is allied generally to work that has begun to emphasise the role of general sensory and motor processes (Buchsbaum & D'Esposito, 2008, 2019; Fuster, 2004; Hickok, 2013). The focus on verbal serial short-term memory and sequence learning was motivated by the research set within the framework that emphasises the role of motor planning in verbal serial recall (e.g. Jones & Macken, 2018) and by the emerging work showing the promise of the approach in relation to Hebb sequence learning (Sjöblom & Hughes, 2020). The framework has begun to be extended to other verbal memory tasks however. For example, the effects of phonological similarity and articulatory suppression on word-form learning in paired-associate tasks have traditionally been taken as evidence for the phonological-store account, particularly the view that the phonological store supports language learning (Baddeley et al., 1998). In the paired-associates task, Sjöblom (2019) used lists of nonwords (paired with known words which were used as the cues at test), manipulating the phonological similarity of the syllables within each nonword or manipulating the phonological similarity of the nonwords to one another. The fact that phonological similarity effects were found (regardless of the particular manipulation) on nonword learning in fact opposes phonological-store models where there should not be an effect of phonological similarity on such long-term sequence learning (Hitch et al., 2009; Page et al., 2006). The effect of within-nonword phonological similarity on learning in particular can be understood in terms of a greater number of speech-planning errors with similar nonwords (Dell, 1984; Shattuck-Hufnagel, 1992). Articulatory suppression also impaired not only paired-associate recall but also paired-

associate learning regardless of presentation modality (visual or auditory). On the perceptual-motor approach, these results again suggest that the motor planning that underpins performance in the short-term task, when the same motor-plan is repeatedly assembled and used, also supports learning. Again, an effect of articulatory suppression also goes against phonological-store based models of sequence learning (Hitch et al., 2009; Page et al., 2006).

The results of the current experiments across both the auditory and visual modalities suggest that perceptual-processing systems interact with regions involved in motor planning. Whilst results of the neuroimaging investigations do not speak to *how* perceptual-motor mapping is achieved, the results do converge generally with behavioural evidence showing the close interplay of perceptual and motor processes. For example, the classic irrelevant sound effect—where task-irrelevant auditory sequences impair serial recall (Colle & Welsh, 1976)—provides evidence for the interplay (in this case interference between) obligatory perceptual and deliberate motor processing. Changing sounds produce order cues as a by-product of automatic auditory-perceptual organisation and are a prerequisite for marked disruption (Jones et al., 1992). These extraneous cues interfere with the active creation of a motor sequence-plan for serial ordering of to-be-remembered items (regardless of their presentation-modality). It was originally supposed that the sound interfered with a passive phonological store rather than with active articulatory rehearsal (Hanley & Bakopoulou, 2003; Hanley & Broadbent, 1987; Salamé & Baddeley, 1982) but evidence has since shown that the irrelevant sound effect is eliminated when articulatory rehearsal is impeded by articulatory suppression and this regardless of whether the to-be-remembered verbal list is presented visually or auditorily (Jones et al., 2004; Hanley & Hayes, 2012). The study of the irrelevant sound effect again points

to a central role for articulatory rehearsal in verbal serial short-term memory. More generally, work within the perceptual-motor framework (including the current research) has sought to reconceptualise the very nature of short-term memory (cf. Jones & Macken, 2018) and has presented a basis for future neuroimaging investigations.

If cognitive neuroscientists, or indeed cognitive psychologists, maintain the view that a temporary storage unit must exist specifically for phonological information, then logic dictates that other stores must exist for other types of input. Such views often fail to address how integrated systems in the brain function, or how far more flexible domain-general operations may support specific types of behaviour. Moreover, the notion of an idle short-term storage unit in the brain violates the computational principle that modules in a hierarchical network all perform a function (Buchsbaum & D'Esposito, 2019). Subsequent to receiving input signals, the function of modules from a neuroscience systems view is to transmit a transformed signal to those at lower or higher levels. By definition, however, the phonological store is proposed to retain information and does not actively modify any received input. It is questionable, therefore, whether assuming that a region in the brain—functionally and anatomically separate to perceptual or motor regions—is a neurophysiologically plausible stance (Buchsbaum & D'Esposito, 2019).

The current thesis has focused on approaching *verbal* serial short-term memory from a perspective that deviates radically from the classic store-based view. However, a recent review of the evidence for the other domain-specific store postulated within the working memory model (Baddeley, 2007; Baddeley & Hitch, 1974)—the visual-spatial sketchpad (specialised for temporarily storing visuo-spatial input)—has called into question the existence of that module too. Morey's (2018) review of the

neuropsychology of short-term memory—which she argued was the main foundation for postulating separate short-term storage modules specialised for different kinds of input— revealed that the case studies comparing performance on verbal, visual, and spatial memory tasks did not solely differ in regard to stimulus-modality: response mode and task-constraints that limited response options were also highlighted. When such differences are controlled across different types of memory task in a healthy sample, variation in performance based on stimulus type is greatly reduced (Ward et al., 2005). Clearly, then, to successfully demonstrate that neuropsychological dissociations between verbal, visual, or spatial memory performance reflects disturbances to specialised short-term stores, more experimental control is necessary across task conditions and procedures. Moreover, the cognitive deficits that some patients exhibited could not easily be explained by recourse to impairments in visual or spatial short-term memory. This would then suggest that any deficits observed in visual-spatial memory may not be the result of specific damage to a visual-spatial store but could be due to a region that is involved in a variety of different types of processing.

The perceptual-motor account also converges with some aspects of other theories of short-term memory such as the *embedded processes model* in which 'short-term memory' is conceptualised as the currently activated portion of long-term memory (e.g., Cowan, 1999, 2016). Both accounts suggest that explaining short-term memory performance does not require the postulation of an isolatable storage unit. However, there are some critical differences between the two accounts: In particular, the view that short-term memory phenomena can be entirely explained in terms of activated nodes in long-term memory is at odds with the perceptual-motor account. Whilst the perceptual-motor account holds that long-term memory undeniably

contributes to short-term memory performance (e.g., Jones & Macken, 2015), additional on-the-fly short-term processing is nevertheless deemed necessary for a comprehensive account (for a similar view but from a phonological-store based perspective, see Norris, 2017, 2019).

In recognition of the need to postulate some form of novel processing in the context of a short-term memory task, Cowan (2019) suggests that such processing takes the form of rapid learning within the long-term memory system, where new associations within a presented stimulus-set are formed. However, it is unclear what supports this rapid learning. While it is widely accepted that the products of learning are represented in a long-term memory system, it is less clear how a long-term memory system itself could learn novel information. It is suggested here, based on the present and other studies, that it is the articulatory planning system that provides the 'short-term' vehicle for binding otherwise unrelated items together as is called for in a serial short-term memory task. Indeed, other authors have recently begun to incorporate the central features of the perceptual-motor account into the embedded processes approach:

"one could...consider activated long-term memories to include fleeting representations temporarily preserved by perceptual systems and information kept active by motor re-instantiation. Sensory-motor recruitment makes it unnecessary to impose dedicated, specialized short-term "slave" systems into the embedded process framework's activated memories: the activation of perceptual and motor systems can serve the memory system without creating redundancy" (Morey et al., 2019, p. 158).

It is argued here, however, in line with the results and theorising in the present thesis, that the motor system does much more than 're-instantiate' representations produced via perceptual systems: it *generates* the temporarily-extended object that will

form the basis of the reproduction of the presented sequence. It is also then the generation (and repeated generation) of a new motor-object that supports the entry of the initially novel input into the long-term memory system (as witnessed, for example, in Hebb sequence learning; Sjöblom & Hughes, 2020).

5.4 Verbal Sequence Learning

The second main aim of the present thesis was to assess whether long-term learning of a repeating verbal sequence (Hebb repetition effect; Hebb 1961) was indicated by a decrease in the BOLD signal throughout the course of each Experiment B in cerebellar lobules HIV-HVI, HVIIB and HVIII. Results from both Experiment 1B (auditory) and Experiment 2B (visual) provided behavioural evidence of sequence learning—where accuracy of the recall of repeating sequences increased whilst that for non-repeating sequences remained static—but the functional imaging results only suggested evidence of general task-set, rather than sequence-specific, learning.

To assess if any excitability decreases occurred, first (linear) and second (quadratic) order polynomial expansions of the repeating-sequence regressor were used to flexibly model any observable changes. Whilst significant results of the *Repeating (Hebb-linear) > Non-repeating (Filler)* contrast revealed activation in several cerebellar lobules (cf. Chapter 3, Table 7 and Chapter 4, Table 12), assessment of the parameter estimate graphs revealed that the difference between the unmodulated and modulated regressors (first and second order polynomial expansions) was in fact what drove the significant results. Parameter estimates in lobules HV, HVI, HVIIA, HVIIB and HVIIIA exhibited the same pattern of activity—a decrease in BOLD-signal—for both unmodulated repeating and non-repeating conditions across both modalities. BOLD-signal decreases specific to the repeating sequences were therefore not observed as predicted but reflected a general learning of how to complete the task.

Speculatively, such results could be interpreted as demonstrating the adaptation of a system geared for motor planning, that is, learning how to prepare the system necessary for assembling a motor plan (based on incoming items) over multiple serial recall trials in succession. The automatisation of the motor planning process is described in Chapter 1, Section 1.6.

As the predicted haemodynamic effects were not observed in either experiment, one must consider the possibility that different haemodynamic markers reflect the acquisition of motor skills. It has been suggested that the decreases in cerebellar activity during progressive motor sequence learning (Grafton et al., 2002; Toni et al, 1994; Tzvi et al, 2014) is related to the processing of prediction errors, which themselves decrease during learning when a motor sequence becomes automatised (Tzvi et al, 2014). A reduction of synaptic efficacy between Purkinje cells and their corresponding inputs are predicted to co-occur with motor learning and Purkinje cell activity is predicted to decline down to 50% below baseline (Gilbert & Thach, 1977; Jirenhed et al., 2007). Alternatively, others suggest that excitability decreases are indicative of to-be-learned stimulus-response associations being processed, as compared to those that are already learned (Wolpert et al., 1998). Experience-dependent adaptive learning processes are a primary feature of cerebellar function (Thach, 1998) but contributions of the cerebellar cortex to motor learning are still a matter of debate. It is also possible that instead, an *increase* in BOLD-signal across the course of the long-term learning experiment may have occurred and could be explored in future investigations of verbal sequence learning. Although many studies have shown excitability decreases in the cerebellum during learning in both motor and cognitive domains (Balsters & Ramnani 2011; Doyon et al. 2002; Imamizu et al. 2000; Penhune & Doyon 2005), other research has shown increases.

Within the domain of verbal short-term memory, Peterburs et al. (2019) hypothesised that activation in lobule HVIII would be sensitive to load (five letters vs. two letters) and thus its activation would be greatest when subjects were required to remember phonologically similar letters in comparison to dissimilar. Using the Sternberg (1966) paradigm but with novel and repeating sequences of letters that were either phonologically similar or dissimilar, they hypothesised that high-load sequences of phonologically similar items would elicit the greatest memory demand and thus activation in lobule HVIIIA would be higher than that of dissimilar letters but, simultaneously, that learning of a repeating similar sequence would elicit a greater decrease in comparison to repeating dissimilar sequences. Decreases in activity were observed in lobule HVI for repeating-dissimilar-high-load trials but not repeating-similar-high-load. Instead, a decrease was observed in lobule HVIIIA for repeating-similar-high-load in comparison to novel-similar-high-load. Based on these results, the authors suggested that lobule HVIIIA may generate association-based predictions of letter sequences that would reduce the likelihood of phonological loop failure before a recall phase. However, the pattern of activations regarding repetition—as well as the known connections of lobule HVI and HVIII—suggest instead that decreases in activation were related to motor planning rather than processes supporting a phonological store. Indeed, results demonstrating modality-specific laterality effects in lobule VIII suggest that the lobule does not function to support a phonological store: Damage to the left lobule HVIII is associated with impaired auditory digit span performance (Chiricozzi et al., 2008; Kirschen et al., 2008; Ravizza et al., 2006). Contrastingly, removal of the right-cerebellar hemisphere (Silveri et al., 1998) results in a phonological similarity effect for auditorily but not visually presented items. Such results therefore indicate that lobule HVIII could not support the operation of a

modality-independent phonological store as disturbances to effects in both modalities should then be observed.

More broadly, learning constant timings of specified finger movement sequences (Jueptner et al., 1997; Ramnani & Passingham, 2001), resulted in increases in cerebellar activity and were attributed to the storage of short-term motor representations into procedural-motor long-term memory. It was argued that as subjects' preparation of responses to timings became automated, the ability to prepare actions increased, thus resulting in increased excitability. In studies where subjects were required to respond to specific timings using unspecified fingers—meaning subjects were unable to prepare action related to a specific effector or series of effectors at a given time—the increases were therefore attributed to learning of timing sequences independently from specific effectors (Sakai et al., 2002b). Other studies have shown different combinations of excitability changes: increased rates of articulation have been associated with increased excitability in the cerebellum and thalamus but a decrease in the striatum (Riecker et al., 2006). Extensive behavioural training in visuo-motor serial reaction time tasks has also shown that activation in the associative striatum observed during the early stages of learning declines to background levels as behaviours become automated (Poldrack et al., 2005) whilst others have shown increases in activation in the sensorimotor striatum when sequences have become automated (Miyachi et al., 2002). These results are consistent with research highlighting the contributions of the basal ganglia to learning and automaticity (Ashby et al., 2010) and suggest that changes indicative of learning could be observed in the network of regions between the cerebellum and cortex and could be investigated in future investigations of verbal sequence learning.

5.5 Limitations and Future Directions

A consideration of possible changes to the current experimental design may offer ideas for future research. As the analyses focussed on activation during the temporal delay during verbal serial rehearsal and whether dynamic changes associated with recalling a repeating sequence would be observed as the requirement for motor planning reduced, changes could be implemented to the experimental design and analyses for future experiments. It should be considered that dynamic changes may occur during the presentation of verbal items (when motor –planning is assumed to begin) and the recall phase. The response mode in the current experiments required subjects to complete a fragment of the initially presented sequence by pressing one of four buttons at a time. Instead, adapting the design to incorporate sparse sampling imaging would allow subjects to produce a vocal response instead. Scans would be acquired following the designated vocal response period but would capture the BOLD-signal of that period given the delay in the onset of the curve. But a vocal response mode may present its own challenges. It would mean, for example, that no visualisable output would be present for subjects to monitor how many responses they have made.

The way in which serial recall data are typically scored constrains interpretations of performance, particularly when using a discrete response mode such as button presses, in a way that may obscure the contribution of dynamic motor processing to that performance (Macken et al., 2015). Whilst using discrete response modes enables simple categorisation of serial recall responses (i.e., correct or incorrect), dynamic 'errors' such as hesitations and self-corrections are typically ignored in memory contexts. And yet these could provide further insight into how motor planning enables verbal serial recall performance and sequence learning. As evidence suggests that the types of articulatory errors that occur in naturalistic speech

occur also in verbal memory tasks (the *error equivalence hypothesis* Ellis, 1980), fine-grained error analysis of continuous (or semi-continuous) vocal output sequences may be particularly informative from a motor-based perspective on verbal serial short-term memory and learning: Measuring how and when subjects self-correct simultaneously with an assessment of whether the frequency or/and type of errors changes as correct motor trajectories become increasingly fluent seems an important next step for the perceptual-motor approach (for further discussion, see Acheson & MacDonald, 2009).

A further consideration for future research designed to adjudicate between the two accounts of short-term memory examined in the present thesis pertains to the reliance on null hypothesis significance testing. The neuroimaging results from the range of contrasts conducted in the present studies clearly demonstrated that a singular region across the parietal lobe was not consistently active across task phases or modalities, thus violating key characteristics of the phonological store concept. It is necessary to consider, however, that the absence of evidence does not necessarily constitute evidence of absence (i.e., evidence for the null hypothesis) and using Bayesian analyses—updating probabilities of hypotheses as more data are collected—could help to further differentiate between the traditional and alternative approach to verbal short-term memory.

It should be noted, however, that the present experimental design did clearly possess the power to detect activation of certain brain regions within each of the task phases and across modalities, meaning that results could indeed confirm or deny the hypotheses drawn from the two accounts. That is, some of the most key conclusions were based on dissociations in the pattern of brain activations across task-phases or modalities rather than simply null effects. A direct statistical comparison between the two models, however, was not viable. Whilst the perceptual-motor approach made

specific predictions about activation in motor planning regions, it did not make specific predictions about what function the parietal lobe may have in verbal serial short-term memory although some attempt at interpretation was made within the empirical chapters. Similarly, whilst proponents of the phonological loop model (e.g., Baddeley, 2003) argue that the phonological store is in principle localisable, there is little or no consensus as to its actual location (Awh et al., 1996; Henson et al., 2000; Paulesu et al., 1993; Salmon et al., 1996; Smith et al., 1996), thus also hindering the formal modelling of the account's predictions.

5.6 Conclusions

The experiments presented in this thesis focussed on re-examining the neural bases of verbal serial-short term memory and learning by recourse to motor planning, perceptual processes, and perceptual-motor processes. Critically, no regions in the parietal lobe (or indeed elsewhere), that had been identified previously as possible substrates of the phonological store were found to be active across the auditory and visual experiments or across presentation and temporal-delay phases. Results from Experiment A across both modalities revealed clearly the common activation of motor planning regions—or those whose output is feeds forward to those regions (e.g. the prefrontal cortex)—in both the cortex and cerebellum and regions specific to the input-modality in question (auditory or visual). Whilst the prediction of BOLD-signal changes in the cerebellum during long-term sequence learning were not confirmed, the results provide the impetus for future studies assessing possible increases in cerebellar activity during motor learning or dynamic changes in activity in sub-cortical structures within the network of regions between the cerebellum and cortex.

REFERENCES

Acheson, D. J., & MacDonald, M. C. (2009). Twisting tongues and memories: Explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language*, *60*(3), 329–350.

Ackermann, H. (2008). Cerebellar contributions to speech production and speech perception: psycholinguistic and neurobiological perspectives. *Trends in Neurosciences*, *31*(6), 265–272.

Ackermann, H., Wildgruber, D., Daum, I., & Grodd, W. (1998). Does the cerebellum contribute to cognitive aspects of speech production? A functional magnetic resonance imaging (fMRI) study in humans. *Neuroscience Letters*, *247*(2–3), 187–190.

Alario, F. X., Chainay, H., Lehericy, S., & Cohen, L. (2006). The role of the supplementary motor area (SMA) in word production. *Brain research*, *1076*(1), 129-143.

Albus, J. S. (1971). A theory of cerebellar function. *Mathematical biosciences*, *10*(1-2), 25-61.

Allport, D. A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of 25 years. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligent, and cognitive neuroscience* (pp. 183–218). Cambridge, MA: MIT Press.

Amunts, K., & Zilles, K. (2015). Architectonic Mapping of the Human Brain beyond Brodmann. *Neuron*, *88*(6), 1086–1107.

Amunts, K., Knoll, A. C., Lippert, T., Pennartz, C. M., Ryvlin, P., Destexhe, A., & Bjaalie, J. G. (2019). The Human Brain Project—Synergy between neuroscience, computing, informatics, and brain-inspired technologies. *PLoS Biology*, *17*(7), e3000344.

Andersen, R. A., & Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annual Review of Neuroscience*, *25*(1), 189-220.

Ashby, F. G., Turner, B. O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in Cognitive Sciences*, *14*(5), 208–215.

Atkins, P. W. B., & Baddeley, A. D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics*, *19*(4), 537–552.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *Psychology of learning and motivation* (Vol. 2, pp. 89-195). Academic Press.

Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppe, R. A., & Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. *Psychological Science*, *7*(1), 25-31.

Babinski, J. (1902). *Sur le role du cervelet dans les actes volitionnels necessitant une succession rapide de mouvements (1)(Diadocoeinesie)*. Masson.

Baddeley, A. D. (1966). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18*(4), 362-365.

Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *302*(1110), 311-324.

Baddeley, A. D. (1986). Working memory. Oxford, England: Clarendon Press.

Baddeley, A. (1988). The role of working memory in vocabulary acquisition. *Bulletin of The Psychonomic Society, 26*(6), 509.

Baddeley, A. (1989). The uses of working memory. In *Memory: interdisciplinary approaches* (pp. 107-123). Springer, New York, NY.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*(10), 829–839.

Baddeley, A. (2007). *Working memory, thought, and action* (Vol. 45). OuP Oxford.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2019). From short-term store to multicomponent working memory: The role of the modal model. *Memory and Cognition*, *47*(4), 575–588.

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*(1), 158-173.

Baddeley, A. D., & Larsen, J. D. (2007). The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative. *Quarterly Journal of Experimental Psychology*, *60*(4), 497–504.

Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology Section A, 36*(2), 233-252.

Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of memory and language*, *27*(5), 586-595.

Badre, D., & D'Esposito, M. (2007). Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex. *Journal of cognitive neuroscience*, *19*(12), 2082-2099.

Badre, D., & D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews Neuroscience*, *10*(9), 659-669.

Baldo, J. V., Klostermann, E. C., & Dronkers, N. F. (2008). It's either a cook or a baker: Patients with conduction aphasia get the gist but lose the trace. *Brain and Language*, *105*(2), 134–140.

Balsters, J. H., Cussans, E., Diedrichsen, J., Phillips, K. A., Preuss, T. M., Rilling, J. K., & Ramnani, N. (2010). Evolution of the cerebellar cortex: The selective expansion of prefrontal-projecting cerebellar lobules. *NeuroImage*, *49*(3), 2045–2052.

Balsters, J. H., & Ramnani, N. (2008). Symbolic representations of action in the human cerebellum. *NeuroImage*, *43*(2), 388–398.

Balsters, J. H., & Ramnani, N. (2011). Cerebellar plasticity and the automation of first-order rules. *Journal of Neuroscience*, *31*(6), 2305–2312.

Balsters, J.H., Whelan, C. D., Robertson, I. H., & Ramnani, N. (2013). Cerebellum and cognition: Evidence for the encoding of higher order rules. *Cerebral Cortex*, *23*(6), 1433–1443.

Barlow, J. S. (2005). *The cerebellum and adaptive control*. Cambridge University Press.

Barton, J. J. S., Fox, C. J., Sekunova, A., & Iaria, G. (2010). Encoding in the visual word form area: An fMRI adaptation study of words versus handwriting. *Journal of Cognitive Neuroscience*, *22*(8), 1649–1661.

Bates, J. F., & Goldman-Rakic, P. S. (1993). Prefrontal connections of medial motor areas in the rhesus monkey. *Journal of Comparative Neurology*, *336*(2), 211-228.

Bauer, L. (2008). Lenition revisited. *Journal of Linguistics*, *44*(1), 605-624.

Beaman, C. P. (2002). Inverting the modality effect in serial recall. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *55*(2), 371–389.

Becker, J. T., MacAndrew, D. K., & Fiez, J. A. (1999). A comment on the functional localization of the phonological storage subsystem of working memory. *Brain and Cognition*, *41*(1), 27-38.

Becker, J. T., Mintun, M. A., Diehl, D. J., Dobkin, J., Martidis, A., Madoff, D. C., & Dekosky, S. T. (1994). Functional neuroanatomy of verbal free recall: A replication study. *Human Brain Mapping*, *1*(4), 284-292.

Benson, D. F., Sheremata, W. A., Bouchard, R., Segarra, J. M., Price, D., & Geschwind, N. (1973). Conduction aphasia: a clinicopathological study. *Archives of Neurology*, *28*(5), 339-346.

Ben-Shachar, M., Dougherty, R. F., Deutsch, G. K., & Wandell, B. A. (2011). The development of cortical sensitivity to visual word forms. *Journal of Cognitive Neuroscience*, *23*(9), 2387–2399.

Bernal, B., & Altman, N. (2010). The connectivity of the superior longitudinal fasciculus: A tractography DTI study. *Magnetic Resonance Imaging*, *28*(2), 217–225.

Besner, D. (1987). Phonology, lexical access in reading, and articulatory suppression: A critical review. *The Quarterly Journal of Experimental Psychology Section A*, *39*(3), 467-478.

Besner, D., Davies, J., & Daniels, S. (1981). Reading for meaning: The effects of concurrent articulation. *The Quarterly Journal of Experimental Psychology*, *33*(4), 415-437.

Binder, J., Frost, J. A., Hammeke, T. A., Bellgowan, P. S. F., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, *10*(5), 512–528.

Binkofski, F., Amunts, K., Stephan, K. M., Posse, S., Schormann, T., Freund, H. J., Zilles, K., & Seitz, R. J. (2000). Broca's region subserves imagery of motion: A combined cytoarchitectonic and fMRI study. *Human Brain Mapping*, *11*(4), 273–285.

Binkofski, F., Buccino, G., Zilles, K., and Fink, G. R. (2004). Supramodal representation of objects and actions in the human inferior temporal and ventral premotor cortex. *Cortex* 40, 159–161.

Bludau, S., Eickhoff, S. B., Mohlberg, H., Caspers, S., Laird, A. R., Fox, P. T., Schleicher, A., Zilles, K., & Amunts, K. (2014). Cytoarchitecture, probability maps and functions of the human frontal pole. *NeuroImage*, *93*, 260–275.

Bokde, A. L., Tagamets, M. A., Friedman, R. B., & Horwitz, B. (2001). Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. *Neuron*, *30*(2), 609-617.

Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *NeuroImage*, *32*(2), 821–841.

Boomer, D. S., & Laver, J. D. (1968). Slips of the tongue. *British Journal of Disorders of Communication*, *3*(1), 2-12.

Boyden, E. S., Katoh, A., & Raymond, J. L. (2004). Cerebellum-dependent learning: The role of multiple plasticity mechanisms. *Annual Review of Neuroscience*, *27*, 581–609.

Braitenberg, V., & Atwood, R. P. (1958). Morphological observations on the cerebellar cortex. *Journal of Comparative Neurology*, *109*(1), 1-33.

Braitenberg, V., Heck, D., Sultan, F., Arbid, M. A., Spoelstra, J., Bjaalie, J. G., & Wessel, K. (1997). The detection and generation of sequences as a key to cerebellar function: experiments and theory. *Behavioral and Brain Sciences*, *20*(2), 229-277.

Bray, S., Arnold, A. E. G. F., Iaria, G., & MacQueen, G. (2013). Structural connectivity of visuotopic intraparietal sulcus. *NeuroImage*, *82*, 137–145.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. MIT press.

Bregman, A. S., & Rudnicky, A. I. (1975). Auditory segregation: Stream or streams?. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(3), 263-267.

Brem, S., Bach, S., Kucian, K., Kujala, J. V., Guttorm, T. K., Martin, E., & Richardson, U. (2010). Brain sensitivity to print emerges when children learn letter–speech sound correspondences. *Proceedings of the National Academy of Sciences*, *107*(17), 7939-7944.

Brendel, B., Hertrich, I., Erb, M., Lindner, A., Riecker, A., Grodd, W., & Ackermann, H. (2010). The contribution of mesiofrontal cortex to the preparation and execution of repetitive syllable productions: An fMRI study. *NeuroImage*, *50*(3), 1219–1230.

Brodal, P. (1979). The pontocerebellar projection in the rhesus monkey: an experimental study with retrograde axonal transport of horseradish peroxidase. *Neuroscience*, *4*(2), 193-208.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth.

Browman, C. P., & Goldstein, L. (1995). Dynamics and articulatory phonology. *Mind as motion*, *175*, 193.

Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. *Cerebral cortex*, *18*(4), 837-845.

Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory - An aggregate analysis of lesion and fMRI data. *Brain and Language*, *119*(3), 119–128.

Buchsbaum, B. R., & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience*, *20*(5), 762–778.

Buchsbaum, B. R., & D'Esposito, M. (2019). A sensorimotor view of verbal working memory. *Cortex*, *112*, 134-148.

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. (2005a). Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron*, *48*(4), 687-697.

Buchsbaum, B. R., Olsen, R. K., Koch, P. F., Kohn, P., Kippenhan, J. S., & Berman, K. F. (2005b). Reading, hearing, and the planum temporale. *Neuroimage*, *24*(2), 444-454.

Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C., & Thomas Yeo, B. T. (2011). The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(5), 2322–2345.

Budisavljevic, S., Dell'Acqua, F., Djordjilovic, V., Miotto, D., Motta, R., & Castiello, U. (2017). The role of the frontal aslant tract and premotor connections in visually guided hand movements. *NeuroImage*, *146*(April 2016), 419–428.

Burani, C., Vallar, G., & Bottini, G. (1991). Articulatory coding and phonological judgements on written words and pictures: The role of the phonological output buffer. *European Journal of Cognitive Psychology*, *3*(4), 379–398.

Burgess, N., & Hitch, G. (2005). Computational models of working memory: Putting long-term memory into context. *Trends in Cognitive Sciences*, *9*(11), 535–541.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*(3), 551–581.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language*, *55*(4), 627–652.

Burton, M. W., Small, S. L., & Blumstein, S. E. (2000). The role of segmentation in phonological processing: An fMRI investigation. *Journal of Cognitive Neuroscience*, *12*(4), 679–690.

Buxton, R. B., & Frank, L. R. (1997). A model for the coupling between cerebral blood flow and oxygen metabolism during neural stimulation. Journal of cerebral blood flow & metabolism, 17 (1), 64–72.

Buxton, R. B., Uludag˘, K., Dubowitz, D. J., & Liu, T. T. (2004). Modelling the hemodynamic response to brain activation. Neuroimage, 23, S220–S233.

Cairo, T. A., Liddle, P. F., Woodward, T. S., & Ngan, E. T. C. (2004). The influence of working memory load on phase specific patterns of cortical activity. *Cognitive Brain Research*, *21*(3), 377–387.

Callan, D., Callan, A., Gamez, M., Sato, M. aki, & Kawato, M. (2010). Premotor cortex mediates perceptual performance. *NeuroImage*, *51*(2), 844–858.

Callan, D. E., Kawato, M., Parsons, L., & Turner, R. (2007). Speech and song: The role of the cerebellum. *Cerebellum*, *6*(4), 321–327.

Carreiras, M., Armstrong, B. C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*(2), 90–98.

Caspers, J., Zilles, K., Eickhoff, S. B., Schleicher, A., Mohlberg, H., & Amunts, K. (2013). Cytoarchitectonical analysis and probabilistic mapping of two extrastriate areas of the human posterior fusiform gyrus. *Brain Structure and Function*, *218*(2), 511-526.

Catani, M., Jones, D. K., & Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology*, *57*(1), 8–16.

Chein, Jason M., & Fiez, J. A. (2001). Dissociation of verbal working memory system components using a delayed serial recall task. *Cerebral Cortex*, *11*(11), 1003–1014.

Chein, J. M., & Fiez, J. A. (2010). Evaluating models of working memory through the effects of concurrent irrelevant information. *Journal of Experimental Psychology: General*, *139*(1), 117.

Chein, J. M., Fissell, K., Jacobs, S., & Fiez, J. A. (2002). Functional heterogeneity within Broca's area during verbal working memory. *Physiology and Behavior*, *77*(4–5), 635–639.

Chein, J. M., Ravizza, S. M., & Fiez, J. A. (2003). Using neuroimaging to evaluate models of working memory and their implications for language processing. *Journal of Neurolinguistics*, *16*(4–5), 315–339.

Chen, S. A., & Desmond, J. E. (2005a). Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. *Neuroimage*, *24*(2), 332-338.

Chen, S. A., & Desmond, J. E. (2005b). Temporal dynamics of cerebro-cerebellar network recruitment during a cognitive task. *Neuropsychologia*, *43*(9), 1227-1237.

Chiricozzi, F. R., Clausi, S., Molinari, M., & Leggio, M. G. (2008). Phonological short-term store impairment after cerebellar lesion: A single case study. *Neuropsychologia*, *46*(7), 1940–1953.

Cholin, J., Levelt, W. J. M., & Schiller, N. O. (2006). Effects of syllable frequency in speech production. *Cognition*, *99*(2), 205–235.

Clower, D. M., West, R. A., Lynch, J. C., & Strick, P. L. (2001). The inferior parietal lobule is the target of output from the superior colliculus, hippocampus, and cerebellum. *Journal of Neuroscience*, *21*(16), 6283-6291.

Coffman, K. A., Dum, R. P., & Strick, P. L. (2011). Cerebellar vermis is a target of projections from the motor areas in the cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(38), 16068–16073.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., & Michel, F. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*(2), 291-307.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, *3*(3), 201-215.

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

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a connectionist approach to behavior and biology in schizophrenia. *Psychological review*, *99*(1), 45.

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. Annual Review of Neuroscience, 22, 319-349.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of verbal learning and verbal behavior*, *15*(1), 17-31.

Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term-memory recall of repeated and novel word lists. *Memory & Cognition*, *21*(4), 539-545.

Condillac, E. B. de. (1971). An essay on the origin of human knowledge(T. Nugent, Tr.). Gainesville, FL: Scholars Facsimiles and Reprints (Originally published 1746)

Conrad, R., & Hull, A. J. (1964). Information, acoustic confusion and memory span. *British journal of psychology*, *55*(4), 429-432.

Conrad, R., & Hull, A. J. (1968). Input modality and the serial position curve in short-term memory. *Psychonomic Science*, *10*(4), 135–136.

Cornelissen, P. L., Kringelbach, M. L., Ellis, A. W., Whitney, C., Holliday, I. E., & Hansen, P. C. (2009). Activation of the left inferior frontal gyrus in the first 200 ms of reading: Evidence from Magnetoencephalography (MEG). *PLoS ONE*, *4*(4).

Coull, J. T., Vidal, F., & Burle, B. (2016). When to act, or not to act: That's the SMA's question. *Current Opinion in Behavioral Sciences*, *8*, 14–21.

Crepel, F., & Jaillard, D. (1991). Pairing of pre-and postsynaptic activities in cerebellar Purkinje cells induces long-term changes in synaptic efficacy in vitro. *The Journal of Physiology*, *432*(1), 123-141.

Crepel, F., Mariani, J., & Delhaye-bouchaud, N. (1976). Evidence for a Multiple Innervation of Purkinje Cells by Climbing Fibers in the Immature Rat Cerebellum. *Journal of Neurobiology, 7*(6), 567–578.

Crosson, B., Sadek, J. R., Maron, L., Gökçay, D., Mohr, C. M., Auerbach, E. J., & Briggs, R. W. (2001). Relative shift in activity from medial to lateral frontal cortex during internally versus externally guided word generation. *Journal of cognitive neuroscience*, *13*(2), 272-283.

Crowder, R. G. (1971). The sound of vowels and consonants in immediate memory. *Journal of Verbal Learning and Verbal Behavior*, *10*(6), 587–596.

D'Angelo, E., & De Zeeuw, C. I. (2009). Timing and plasticity in the cerebellum: focus on the granular layer. *Trends in Neurosciences*, *32*(1), 30–40.

Damasio, H., & Damasio, A. R. (1980). The anatomical basis of conduction aphasia. *Brain*, *103*(2), 337-350.

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3773–3800.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*(2), 384–398.

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262.

Démonet, J. F., Fiez, J. A., Paulesu, E., Petersen, S. E., & Zatorre, R. J. (1996). PET studies of phonological processing: A critical reply to Poeppel.

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human

cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980.

Desmond, J. E. (2001). Cerebellar involvement in cognitive function: evidence from neuroimaging. *International Review of Psychiatry*, *13*(4), 283-294.

Desmond, J. E., Chen, S. A., & Shieh, P. B. (2005). Cerebellar transcranial magnetic stimulation impairs verbal working memory. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, *58*(4), 553-560.

Desmond, J. E., Gabrieli, J. D. E., Wagner, A. D., Ginier, B. L., & Glover, G. H. (1997). *Lobular Patterns of Cerebellar Activation in Verbal Working- Memory and Finger-Tapping Tasks as Revealed by Functional MRI*. *17*(24), 9675–9685.

Devlin, J. T., Matthews, P. M., & Rushworth, M. F. (2003). Semantic processing in the left inferior prefrontal cortex: a combined functional magnetic resonance imaging and transcranial magnetic stimulation study. *Journal of cognitive neuroscience*, *15*(1), 71-84.

Diedrichsen, J. (2006). A spatially unbiased atlas template of the human cerebellum. *NeuroImage*, *33*(1), 127–138.

Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., & Ramnani, N. (2009). A probabilistic MR atlas of the human cerebellum. *NeuroImage*, *46*(1), 39–46.

Diedrichsen, J., Maderwald, S., Küper, M., Thürling, M., Rabe, K., Gizewski, E. R., & Timmann, D. (2011). Imaging the deep cerebellar nuclei: a probabilistic atlas and normalization procedure. *Neuroimage*, *54*(3), 1786-1794.

Diedrichsen, J., & Zotow, E. (2015). Surface-based display of volume-averaged cerebellar imaging data. *PLoS ONE*, *10*(7), 1–18.

Doyon, J., & Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Current opinion in neurobiology*, *15*(2), 161-167.

Doyon, J., Penhune, V., & Ungerleider, L. G. (2003). Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia*, *41*(3), 252–262.

Doyon, J., Song, A. W., Karni, A., Lalonde, F., Adams, M. M., & Ungerleider, L. G. (2002). Experience-dependent changes in cerebellar contributions to motor sequence learning. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(2), 1017–1022.

Duann, J. R., Jung, T. P., Kuo, W. J., Yeh, T. C., Makeig, S., Hsieh, J. C., et al. (2002). Single-trial variability in event-related BOLD signals. *Neuroimage* 15, 823–835.

Dum, R. P., & Strick, P. L. (2003). An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. *Journal of neurophysiology*, *89*(1), 634-639.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature reviews neuroscience*, *2*(11), 820-829.

Durisko, C., & Fiez, J. A. (2010). Functional activation in the cerebellum during working memory and simple speech tasks. *Cortex*, *46*(7), 896–906.

Duvernoy H.M., & Bourgouin, P. (1999). *The human brain: surface, three-dimensional sectional anatomy and MRI.* Vienna: Springer.

Ellis, A. W. (1980). Errors in speech and short-term memory: The effects of phonemic similarity and syllable position. *Journal of Verbal Learning and Verbal Behavior*, *19*(5), 624–634.

Fegen, D., Buchsbaum, B. R., & Esposito, M. D. (2015). NeuroImage The effect of rehearsal rate and memory load on verbal working memory. *NeuroImage*, *105*,

120–131.

Fiez, J. A., Raife, E. A., Petersen, S. E., Balota, D. A., & Raichle, E. (1996). *A Positron Emission Tomography Maintenance of Verbal Information Study of the Short-Term*. *76*(2), 808–822.

Fox, C. A., & Barnard, J. W. (1957). A quantitative study of the Purkinje cell dendritic branchlets and their relationship to afferent fibres. *Journal of Anatomy, 91*(3), 299–313.

Flowers, D. L., Jones, K., Noble, K., VanMeter, J., Zeffiro, T. A., Wood, F. B., & Eden, G. F. (2004). Attention to single letters activates left extrastriate cortex. *Neuroimage*, *21*(3), 829-839.

Frazier, J. A., Chiu, S., Breeze, J. L., Makris, N., Lange, N., Kennedy, D. N., Herbert, M. R., Bent, E. K., Koneru, V. K., Dieterich, M. E., Hodge, S. M., Rauch, S. L., Grant, P. E., Cohen, B. M., Seidman, L. J., Caviness, V. S., & Biederman, J. (2005). Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. *American Journal of Psychiatry*, *162*(7), 1256–1265.

Friston, K. J., Frith, C. D., Frackowiak, R. S., & Turner, R. (1995a). Characterizing dynamic brain responses with fMRI: a multivariate approach. *Neuroimage*, *2*(2), 166-172.

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

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R. 1998. Event-related fMRI: characterizing differential responses. Neuroimage 7:30-40.

Funahashi, S. (2001). Neuronal mechanisms of executive control by the prefrontal cortex. *Neuroscience research*, *39*(2), 147-165.

Fuster, J. M. (1997). Network memory. *Trends in neurosciences*, *20*(10), 451-459.

Fuster, J. M. (2000). Executive frontal functions. *Experimental brain research*, *133*(1), 66-70.

Fuster, J. M. (2004). Upper processing stages of the perception–action cycle. *Trends in cognitive sciences*, *8*(4), 143-145.

Galletti, C., Gamberini, M., Kutz, D. F., Fattori, P., Luppino, G., & Matelli, M. (2001). The cortical connections of area V6: An occipito-parietal network processing visual information. *European Journal of Neuroscience*, *13*(8), 1572–1588.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513–543.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory*, *2*(2), 103-127.

Gilbert, P. F. C., & Thach, W. T. (1977). Purkinje cell activity during motor learning. *Brain research*, *128*(2), 309-328.

Giraud, A. L., Kell, C., Thierfelder, C., Sterzer, P., Russ, M. O., Preibisch, C., & Kleinschmidt, A. (2004). Contributions of Sensory Input, Auditory Search and Verbal Comprehension to Cortical Activity during Speech Processing. *Cerebral Cortex*, *14*(3), 247–255.

Goldstein, J. M., Seidman, L. J., Makris, N., Ahern, T., O'Brien, L. M., Caviness, V. S., Kennedy, D. N., Faraone, S. V., & Tsuang, M. T. (2007). Hypothalamic Abnormalities in Schizophrenia: Sex Effects and Genetic Vulnerability.

Biological Psychiatry, *61*(8), 935–945.

Goodglass, H. (1992). Conduction aphasia. *Lawrence Earlbaum Associate*

Gow, D. W. (2012). The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain and Language*, *121*(3), 273–288.

Goodale, M. A., & Milner, A. D. (2018). Two visual pathways—Where have they taken us and where will they lead in future?. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, *15*(1), 20-25.

Grafton, Scot T., Woods, R. P., & Tyszka, M. (1994). Functional imaging of procedural motor learning: Relating cerebral blood flow with individual subject performance. *Human Brain Mapping*, *1*(3), 221–234.

Grafton, Scott T., Hazeltine, E., & Ivry, R. B. (2002). Motor sequence learning with the nondominant left hand: A PET functional imaging study. *Experimental Brain Research*, *146*(3), 369–378.

Grasby, P. M., Frith, C. D., Friston, K. J., Bench, C. R. S. F., Frackowiak, R. S. J., & Dolan, R. J. (1993). Functional mapping of brain areas implicated in auditory—verbal memory function. *Brain*, *116*(1), 1-20.

Graves, W. W., Grabowski, T. J., Mehta, S., & Gupta, P. (2008). The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. *Journal of Cognitive Neuroscience*, *20*(9), 1698–1710.

Green, E., & Howes, D. H. (1977). The nature of conduction aphasia: A study of anatomic and clinical features and of underlying mechanisms. In *Studies in neurolinguistics* (pp. 123-156). Academic Press.

Grenfell-Essam, R., Ward, G., & Tan, L. (2017). Common modality effects in immediate free recall and immediate serial recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, *43*(12), 1909–1933.

Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of communication disorders*, *39*(5), 350-365.

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and language*, *96*(3), 280-301.

Gupta, P. (2005). Primacy and recency in nonword repetition. *Memory*, *13*(3–4), 318–324.

Greene, R. L. (1991). Serial recall of two-voice lists: Implications for theories of auditory recency and suffix effects. Memory & Cognition, 19(1), 72-78.

Grill-Spector, K. (2003). The neural basis of object perception. *Current opinion in neurobiology*, *13*(2), 159-166.

Hanley, J. R., & Bakopoulou, E. (2003). Irrelevant speech, articulatory suppression and phonological similarity: A test of the phonological loop model and the feature model. *Psychonomic Bulletin and Review*, *10*(2), 435–444.

Hanley, J. R., & Broadbent, C. (1987). The effect of unattended speech on serial recall following auditory presentation. *British Journal of Psychology*, *78*(3), 287–297.

Hanley, J. R., & Hayes, A. (2012). The irrelevant sound effect under articulatory suppression: Is it a suffix effect? *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(2), 482–487.

Hayter, A. L., Langdon, D. W., & Ramnani, N. (2007). Cerebellar contributions to working memory. *NeuroImage*, *36*(3), 943–954.

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. *Brain mechanisms and learning*, *37*, 46.

Heim, S., Eickhoff, S. B., & Amunts, K. (2009). Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. *NeuroImage*, *48*(3), 616–624.

Henson, R. N. A., Burgess, N., & Frith, C. D. (2000). Recoding, storage, rehearsal and grouping in verbal short-term memory: An fMRI study. *Neuropsychologia*, *38*(4), 426–440.

Henson R, Rugg MD, Friston K. 2001. The choice of basis function in event-related fMRI. Neuroimage 9:125.

Herman, A. B., Houde, J. F., Vinogradov, S., & Nagarajan, S. S. (2013). Parsing the phonological loop: Activation timing in the dorsal speech stream determines accuracy in speech reproduction. *Journal of Neuroscience*, *33*(13), 5439–5453.

Hesling, I., Dilharreguy, B., Clément, S., Bordessoules, M., & Allard, M. (2005). Cerebral mechanisms of prosodic sensory integration using low-frequency bands of connected speech. *Human Brain Mapping*, *26*(3), 157–169.

Hewes, G. W. (1973). An explicit formulation of the relationship between tool-using, tool-making, and the emergence of language. *Visible Language*, *7*(2), 101-127.

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

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature reviews neuroscience*, *13*(2), 135-145.

Hickok, G. (2013). The architecture of speech production and the role of the phoneme in speech processing. *Language, Cognition and Neuroscience*, *29*(1), 2-20.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory–Motor Interaction Revealed by fMRI: Speech, Music, and Working Memory in Area Spt. *Journal of Cognitive Neuroscience*, *15*(5), 673–682.

Hickok, G., Okada, K., & Serences, J. T. (2009). Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *Journal of Neurophysiology*, *101*(5), 2725–2732.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in cognitive sciences*, *4*(4), 131-138.

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

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, *8*(5), 393-402.

Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current opinion in neurobiology*, *12*(2), 217-222.

Hikosaka, O., Sakai, K., Miyauchi, S., Takino, R., Sasaki, Y., & Pütz, B. (1996). Activation of human presupplementary motor area in learning of sequential procedures: A functional MRI study. *Journal of Neurophysiology*, *76*(1), 617–621.

Hitch, G. J., Flude, B., & Burgess, N. (2009). Slave to the rhythm: Experimental tests of a model for verbal short-term memory and long-term sequence learning. *Journal of Memory and Language*, *61*(1), 97-111.

Holmes, G. (1939). The cerebellum of man. *Brain*, *62*(1), 1-30.

Hommel, B. (2010). Grounding attention in action control: The intentional control of selection. *Effortless attention: A new perspective in the cognitive science of attention and action*, 121-140.

Hoshi, E. (2008). Differential involvement of the prefrontal, premotor, and primary motor cortices in rule-based motor behavior. *Neuroscience of rule-guided behavior*, 159-175.

Houghton, G., & Tipper, S. P. (1996). Inhibitory mechanisms of neural and cognitive control: Applications to selective attention and sequential action. *Brain and Cognition*, *30*(1), 20–43.

Hsieh, L., Gandour, J., Wong, D., & Hutchins, G. D. (2001). Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain and Language*, *76*(3), 227–252.

Huerta, M. F., Krubitzer, L. A., & Kaas, J. H. (1986). Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys: I. Subcortical connections. *Journal of Comparative Neurology*, *253*(4), 415-439.

Huerta, M. F., Krubitzer, L. A., & Kaas, J. H. (1987). Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys II. Cortical connections. *Journal of Comparative Neurology*, *265*(3), 332-361.

Hughes, R. W., Chamberland, C., Tremblay, S., & Jones, D. M. (2016). Perceptual-motor determinants of auditory-verbal serial short-term memory. *Journal of Memory and Language*, *90*, 126–146.

Hughes, R. W., Harvey, H., & Mills, J. L. (2021). *Nonword Repetition: The Role of Articulatory Planning,* Manuscript in preparation.

Hughes, R. W., & Marsh, J. E. (2017). The functional determinants of short-term memory: Evidence from perceptual-motor interference in verbal serial recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, *43*(4), 537–551.

Hughes, R. W., Marsh, J. E., & Jones, D. M. (2009). Perceptual-Gestural (Mis)Mapping in Serial Short-Term Memory: The Impact of Talker Variability. *Journal of Experimental Psychology: Learning Memory and Cognition*, *35*(6), 1411–1425.

Hughes, R. W., Marsh, J. E., & Jones, D. M. (2011). Role of serial order in the impact of talker variability on short-term memory: Testing a perceptual organization-based account. *Memory and Cognition*, *39*(8), 1435–1447.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101-144.

Indefrey, P., Kleinschmidt, A., Merboldt, K. D., Krüger, G., Brown, C., Hagoort, P., & Frahm, J. (1997). Equivalent responses to lexical and nonlexical visual stimuli in occipital cortex: A functional magnetic resonance imaging study. *Neuroimage*, *5*(1), 78-81.

Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., PuÈtz, B., & Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, *403*(6766), 192-195.

Ito, M. (1993). Movement and thought: identical control mechanisms by the cerebellum.

Ito, M. (2000). Mechanisms of motor learning in the cerebellum. *Brain Research*, *886*(1–2), 237–245.

Ito, M. (2002). Historical review of the significance of the cerebellum and the role of Purkinje cells in motor learning. *Annals of the New York Academy of Sciences*, *978*(1), 273-288.

Ito, M. (2005). Bases and implications of learning in the cerebellum—adaptive control and internal model mechanism. *Progress in brain research*, *148*, 95-109.

Jacquemot, C., Pallier, C., LeBihan, D., Dehaene, S., & Dupoux, E. (2003). Phonological Grammar Shapes the Auditory Cortex: A Functional Magnetic Resonance Imaging Study. *Journal of Neuroscience*, *23*(29), 9541–9546.

Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences*, *10*(11), 480–486.

Jirenhed, D. A., Bengtsson, F., & Hesslow, G. (2007). Acquisition, extinction, and reacquisition of a cerebellar cortical memory trace. *Journal of Neuroscience*, *27*(10), 2493–2502.

Jobard, G., Crivello, F., & Tzourio-Mazoyer, N. (2003). Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *NeuroImage*, *20*(2), 693–712.

Johansen-Berg, H., Behrens, T. E. J., Robson, M. D., Drobnjak, I., Rushworth, M. F. S., Brady, J. M., & Matthews, P. M. (2004). Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proceedings of the National Academy of Sciences*, *101*(36), 13335-13340.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2006). Perceptual organization masquerading as phonological storage: Further support for a perceptual-gestural view of short-term memory. *Journal of Memory and language*, *54*(2), 265-281.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2007). The phonological store abandoned. *Quarterly Journal of Experimental Psychology*, *60*(4), 505–511.

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. Cognition, 144, 1-13.

Jones, D. M., & Macken, B. (2018). In the Beginning Was the Deed: Verbal Short-Term Memory as Object-Oriented Action. *Current Directions in Psychological*

Science, *27*(5), 351–356.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*(3), 656–674.

Jones, D., Madden, C., & Miles, C. (1992). Privileged Access by Irrelevant Speech to Short-term Memory: The Role of Changing State. *The Quarterly Journal of Experimental Psychology Section A*, *44*(4), 645–669.

Jonides, John, Eric H. Schumacher, Edward E. Smith, Erick J. Lauber, Edward Awh, Satoshi Minoshima, and Robert A. Koeppe. "Verbal working memory load affects regional brain activation as measured by PET." *Journal of cognitive neuroscience* 9, no. 4 (1997): 462-475.

Jueptner, M., Stephan, K. M., Frith, C. D., Brooks, D. J., Frackowiak, R. S., & Passingham, R. E. (1997). Anatomy of motor learning. I. Frontal cortex and attention to action. *Journal of neurophysiology*, *77*(3), 1313-1324.

Kawashima, R., Okuda, J., Umetsu, A., Sugiura, M., Inoue, K., Suzuki, K., Tabuchi, M., Tsukiura, T., Narayan, S. L., Nagasaka, T., Yanagawa, I., Fujii, T., Takahashi, S., Fukuda, H., & Yamadori, A. (2000). Human cerebellum plays an important role in memory-timed finger movement: An fMRI study. *Journal of Neurophysiology*, *83*(2), 1079–1087.

Kawato, M., & Gomi, H. (1992). A computational model of four regions of the cerebellum based on feedback-error learning. *Biological cybernetics*, *68*(2), 95-103.

Kawato, M., & Wolpert, D. (1998, January). Internal models for motor control. In *Novartis Foundation Symposium* (pp. 291-303). Wiley.

Kelly, R. M., & Strick, P. L. (2003). Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *Journal of Neuroscience*, *23*(23), 8432–8444.

Kemeny, S., Xu, J., Park, G. H., Hosey, L. A., Wettig, C. M., & Braun, A. R. (2006). Temporal dissociation of early lexical access and articulation using a delayed naming task - An fMRI study. *Cerebral Cortex*, *16*(4), 587–595.

Kim, S. G., Rostrup, E., Larsson, H. B., Ogawa, S., & Paulson, O. B. (1999). Determination of relative CMRO2 from CBF and BOLD changes: significant increase of oxygen consumption rate during visual stimulation. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *41*(6), 1152-1161.

Kirschen, M. P., Chen, S. H. A., & Desmond, J. E. (2010). Modality specific cerebro-cerebellar activations in verbal working memory: An fMRI study. *Behavioural Neurology*, *23*(1–2), 51–63.

Kirschen, M. P., Chen, S. H. A., Schraedley-Desmond, P., & Desmond, J. E. (2005). Load- and practice-dependent increases in cerebro-cerebellar activation in verbal working memory: An fMRI study. *NeuroImage*, *24*(2), 462–472.

Kirschen, M. P., Davis-ratner, M. S., Milner, M. W., & Chen, S. H. A. (2008). *Verbal memory impairments in children after cerebellar tumor resection*. *20*, 39–53.

Klein, I., Paradis, A. L., Poline, J. B., Kosslyn, S. M., & Le Bihan, D. (2000). Transient activity in the human calcarine cortex during visual-mental imagery: An event-related fMRI study. *Journal of Cognitive Neuroscience*, *12*(SUPPL. 2), 15–23.

Knolle, F., Schröger, E., & Kotz, S. A. (2013). Cerebellar contribution to the prediction of self-initiated sounds. *Cortex*, *49*(9), 2449–2461.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping*, *30*(3), 859–873.

Koenigs, M., Acheson, D. J., Barbey, A. K., Solomon, J., Postle, B. R., & Grafman, J.

(2011). Areas of left perisylvian cortex mediate auditory-verbal short-term memory. *Neuropsychologia*,*49*(13),3612–3619.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature reviews neuroscience*, *2*(9), 635-642.

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends in cognitive sciences*, *14*(9), 392-399.

Krienen, F. M., & Buckner, R. L. (2009). Segregated fronto-cerebellar circuits revealed by intrinsic functional connectivity. *Cerebral cortex*, *19*(10), 2485-2497.

Krishnan, S., Alcock, K. J., Mercure, E., Leech, R., Barker, E., Karmiloff-Smith, A., & Dick, F. (2013). Articulating novel words: Children's oromotor skills predict nonword repetition abilities.

Krishnan, S., Alcock, K. J., Carey, D., Bergström, L., Karmiloff-Smith, A., & Dick, F. (2017). Fractionating nonword repetition: The contributions of short-term memory and oromotor praxis are different. *Plos one*, *12*(7), e0178356.

Kujovic, M., Zilles, K., Malikovic, A., Schleicher, A., Mohlberg, H., Rottschy, C., Eickhoff, S. B., & Amunts, K. (2013). Cytoarchitectonic mapping of the human dorsal extrastriate cortex. *Brain Structure and Function*, *218*(1), 157–172.

Kuypers, H. G., Szwarcbart, M. K., Mishkin, M., & Rosvold, H. E. (1965). Occipitotemporal corticocortical connections in the rhesus monkey. *Experimental neurology*, *11*(2), 245-262.

Larsell, O., & Jansen, J. (1972). *The comparative anatomy and histology of the cerebellum: the human cerebellum, cerebellar connections, and cerebellar cortex* (Vol. 3). University of Minnesota Press.

Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *Journal of neuroscience*, *26*(51), 13128-13142.

Lashley, K. (1951). The problem of serial order in psychology. *Cerebral mechanisms in behavior. New York: Wiley*.

Leff, A. P., Schofield, T. M., Crinion, J. T., Seghier, M. L., Grogan, A., Green, D. W., & Price, C. J. (2009). *substrate for auditory short-term memory and with stroke*.

Leggio, M., & Molinari, M. (2015). Cerebellar Sequencing: a Trick for Predicting the Future. *Cerebellum*, *14*(1), 35–38.

Lehéricy, S., Benali, H., Van De Moortele, P.-F., Pélégrini-issac, M., Waechter, T., Ugurbil, K., & Doyon, J. (2005). Distinct basal ganglia territories are engaged in early and advanced motor sequence learning. *Pnas*, *102*(35), 12566–12571.

Levelt, W. J., (1989). *Speaking: From intention to articulation.-" A bradford book."*. MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75.

Levelt, W. J., & Wheeldon, L. (1994). Do speakers have access to a mental syllabary?. *Cognition*, *50*(1-3), 239-269.

Leybaert, J., & Lechat, J. (2001). Phonological similarity effects in memory for serial order of cued speech.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1986). Does the cerebellum contribute to mental skills?. *Behavioral neuroscience*, *100*(4), 443.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, *74*(6), 431.

Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. *Nature, 453,* 869-878.

Lovett, M. C., Reder, L. M., & Lebiere, C. (1999). Modeling working memory in a unified architecture: An ACT-R perspective.

Lu, X., Miyachi, S., Ito, Y., Nambu, A., & Takada, M. (2007). Topographic distribution of output neurons in cerebellar nuclei and cortex to somatotopic map of primary motor cortex. *European journal of Neuroscience*, *25*(8), 2374-2382.

Lu, M. -T, Preston, J. B., & Strick, P. L. (1994). Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *Journal of Comparative Neurology*, *341*(3), 375–392.

Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1993). Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *Journal of Comparative Neurology*, *338*(1), 114–140.

Mackay, D. G. (1970). *Spoonerisms : the Structure of Errors*. *8*.

Macken, B., Taylor, J. C., Kozlov, M. D., Hughes, R. W., & Jones, D. M. (2016). Memory as embodiment: The case of modality and serial short-term memory. *Cognition*, *155*, 113–124.

Macuga, K. L., & Frey, S. H. (2012). Neural representations involved in observed, imagined, and imitated actions are dissociable and hierarchically organized. *NeuroImage*, *59*(3), 2798–2807.

Makris, N., Goldstein, J. M., Kennedy, D., Hodge, S. M., Caviness, V. S., Faraone, S. V., Tsuang, M. T., & Seidman, L. J. (2006). Decreased volume of left and total anterior insular lobule in schizophrenia. *Schizophrenia Research*, *83*(2–3), 155–171.

Malikovic, A., Amunts, K., Schleicher, A., Mohlberg, H., Kujovic, M., Palomero-Gallagher, N., ... & Zilles, K. (2016). Cytoarchitecture of the human lateral occipital cortex: mapping of two extrastriate areas hOc4la and hOc4lp. *Brain Structure and Function*, *221*(4), 1877-1897.

Manni, E., & Petrosini, L. (2004). A century of cerebellar somatotopy: a debated representation. *Nature Reviews Neuroscience*, *5*(3), 241-249.

Manto, M., Bower, J. M., Conforto, A. B., Delgado-García, J. M., Da Guarda, S. N. F., Gerwig, M.,& Timmann, D. (2012). Consensus paper: roles of the cerebellum in motor control—the diversity of ideas on cerebellar involvement in movement. *The Cerebellum*, *11*(2), 457-487.

Marr, D. (1969). A theory of cerebellar cortex. *The Journal of Physiology*, *202*(2), 437–70.

Marr, D., & Thach, W. T. (1991). A theory of cerebellar cortex. In *From the Retina to the Neocortex* (pp. 11-50). Birkhäuser Boston.

Mars, R. B., Jbabdi, S., Sallet, J., O'Reilly, J. X., Croxson, P. L., Olivier, E., Noonan, M. A. P., Bergmann, C., Mitchell, A. S., Baxter, M. G., Behrens, T. E. J., Johansen-Berg, H., Tomassini, V., Miller, K. L., & Rushworth, M. F. S. (2011). Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. *Journal of Neuroscience*, *31*(11), 4087–4100.

Marvel, C. L., & Desmond, J. E. (2010a). Functional topography of the cerebellum in verbal working memory. *Neuropsychology Review*, *20*(3), 271–279.

Marvel, C. L., & Desmond, J. E. (2010b). The contributions of cerebro-cerebellar circuitry to executive verbal working memory. *Cortex*, *46*(7), 880–895.

Massaro, D. W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review*, *79*(2), 124–145.

Maurer, U., Brem, S., Kranz, F., Bucher, K., Benz, R., Halder, P., Steinhausen, H. C., & Brandeis, D. (2006). Coarse neural tuning for print peaks when children learn to read. *NeuroImage*, *33*(2), 749–758.

McGettigan, C., Warren, J. E., Eisner, F., Marshall, C. R., Shanmugalingam, P., & Scott, S. K. (2011). Neural correlates of sublexical processing in phonological working memory. *Journal of cognitive neuroscience*, *23*(4), 961-977.

McKelvie, S. J. (1987). Learning and awareness in the Hebb digits task. *The Journal of General Psychology*, *114*(1), 75-88.

Miyachi, S., Hikosaka, O., & Lu, X. (2002). Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Experimental brain research*, *146*(1), 122-126.

Mei, L., Xue, G., Chen, C., Xue, F., Zhang, M., & Dong, Q. (2010). The "visual word form area" is involved in successful memory encoding of both words and faces. *NeuroImage*, *52*(1), 371–378.

Meister, I. G., Buelte, D., Staedtgen, M., Boroojerdi, B., & Sparing, R. (2009). The dorsal premotor cortex orchestrates concurrent speech and fingertapping movements. *European Journal of Neuroscience*, *29*(10), 2074–2082.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *2*(1), 1–21.

Meyer, M., Zysset, S., Von Cramon, D. Y., & Alter, K. (2005). Distinct fMRI responses to laughter, speech, and sounds along the human peri-sylvian cortex. *Cognitive Brain Research*, *24*(2), 291–306.

Middleton, F. A., & Strick, P. L. (2001). Cerebellar projections to the prefrontal cortex of the primate. *Journal of neuroscience*, *21*(2), 700-712.

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural networks*, *9*(8), 1265-1279.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, *63*(2), 81.

Miller, E. K., & Cohen, J. D. (2001). An integrate theory of PFC function. *Annual Review of Neuroscience*, *24*, 167–202.

Miller, G. A., Eugene, G., & Pribram, K. H. (2017). *Plans and the Structure of Behaviour* (pp. 369-382). Routledge.

Mills, J. L. (2018). *Skill Learning: Brain Systems, Eye Movements, Ageing, and Driver Behaviour* (Doctoral dissertation, Royal Holloway, University of London).

Mills, J.L., Argyropoulos G., Ramnani, N. (*under review*). Memory-guided Oculomotor Sequences: Prefrontal and Cerebellar Working Memory mechanisms.

Morey, C. C. (2018). The case against specialized visual-spatial short-term memory. *Psychological Bulletin*, *144*(8), 849.

Molinari, M., Chiricozzi, F. R., Clausi, S., Tedesco, A. M., De Lisa, M., & Leggio, M. G. (2008). Cerebellum and detection of sequences, from perception to cognition. *Cerebellum*, *7*(4), 611–615.

Morey, C. C., Rhodes, S., & Cowan, N. (2019). Sensory-motor integration and brain lesions: Progress toward explaining domain-specific phenomena within domain-general working memory. *Cortex*, *112*, 149–161.

Mosse, E. K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the

acquisition of phonological forms in children. *Quarterly Journal of Experimental Psychology*, *61*(4), 505–514.

Mostofsky, S. H., Schafer, J. G. B., Abrams, M. T., Goldberg, M. C., Flower, A. A., Boyce, A., Courtney, S. M., Calhoun, V. D., Kraut, M. A., Denckla, M. B., & Pekar, J. J. (2003). fMRI evidence that the neural basis of response inhibition is task-dependent. *Cognitive Brain Research*, *17*(2), 419–430.

Mottolese, C., Richard, N., Harquel, S., Szathmari, A., Sirigu, A., & Desmurget, M. (2013). Mapping motor representations in the human cerebellum. *Brain*, *136*(1), 330–342.

Murdoch, B. E. (2010). The cerebellum and language: historical perspective and review. *Cortex*, *46*(7), 858-868.

Murray, D. J. (1968). Articulation and acoustic confusability in short-term memory. *Journal of Experimental Psychology*, *78*(4p1), 679.

Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience*, *9*(11), 856–869.

Neumann, O. (1996). Theories of attention. In O. Neumann & A. F. Sanders (Eds.). Handbook of perception and action (Vol. 3, pp. 389–446). London, England: Academic Press.

Neumann, O. (1987). Beyond capacity: A functional view of attention. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action.* Hillsdale, NJ: Erlbaum.

Ng, H. B. T., Kao, K. L. C., Chan, Y. C., Chew, E., Chuang, K. H., & Chen, S. H. A. (2016). Modality specificity in the cerebro-cerebellar neurocircuitry during working memory. *Behavioural Brain Research*, *305*, 164–173.

Nicholls, A. P., & Jones, D. M. (2002). Capturing the Suffix: Cognitive Streaming in

Immediate Serial Recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, *28*(1), 12–28.

Norman, D. A., & Rumelhart, D. E. (1970). A system for perception and memory. In *Models of human memory* (pp. 19-64). Academic Press New York.

Norris, D. (2017). Short-Term Memory and Long-Term Memory are Still Different. Psychological Bulletin, 143(9), 992-1009.

Norris, D. (2019). Even an activated long-term memory system still needs a separate short-term store: A reply to Cowan (2019). Psychological Bulletin, 145(8), 848-853.

Norris, D., Page, M. P., & Hall, J. (2018). Learning nonwords: The Hebb repetition effect as a model of word learning. *Memory*, *26*(6), 852-857.

Nooteboom, S. G. (1967). Some regularities in phonemic speech errors. *IPO Annual progress report*, *2*, 65-70.

O'reilly, J. X., Beckmann, C. F., Tomassini, V., Ramnani, N., & Johansen-Berg, H. (2010). Distinct and overlapping functional zones in the cerebellum defined by resting state functional connectivity. *Cerebral cortex*, *20*(4), 953-965.

Okada, K., Matchin, W., & Hickok, G. (2018). Neural evidence for predictive coding in auditory cortex during speech production. *Psychonomic Bulletin and Review*, *25*(1), 423–430.

Orioli, P. J., & Strick, P. L. (1989). Cerebellar connections with the motor cortex and the arcuate premotor area: An analysis employing retrograde transneuronal transport of WGA-HRP. *Journal of Comparative Neurology*, *288*(4), 612–626.

Pa, J., & Hickok, G. (2008). A parietal-temporal sensory-motor integration area for the human vocal tract: Evidence from an fMRI study of skilled musicians.

Neuropsychologia, *46*(1), 362–368.

Page, Mike P.A., Cumming, N., Norris, D., Hitch, G. J., & McNeil, A. M. (2006). Repetition learning in the immediate serial recall of visual and auditory materials. *Journal of Experimental Psychology: Learning Memory and Cognition*, *32*(4), 716–733.

Page, M. P.A., & Norris, D. (2009a). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3737–3753.

Page, M. P., & Norris, D. (2009b). Is there a common mechanism underlying word-form learning and the Hebb repetition effect? Experimental data and a modelling framework. *Interactions between short-term and long-term memory in the verbal domain*, 136-156.

Page, Mike P.A., Madge, A., Cumming, N., & Norris, D. G. (2007). Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus. *Journal of Memory and Language*, *56*(1), 49–64.

Palmer, E. D., Rosen, H. J., Ojemann, J. G., Buckner, R. L., Kelley, W. M., & Petersen, S. E. (2001). An event-related fMRI study of overt and covert word stem completion. *Neuroimage*, *14*(1), 182-193.

Pandya, D. N., & Kuypers, H. G. (1968). Cortico-cortical connections in the rhesus monkey. *Brain research*, *13*(1), 13-36.

Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., & Tzourio-Mazoyer, N. (2000). A common language network for comprehension and production: a contribution to the definition of language epicenters with PET. *Neuroimage*, *11*(4), 347-357.

Papoutsi, M., De Zwart, J. A., Jansma, J. M., Pickering, M. J., Bednar, J. A., & Horwitz, B. (2009). From phonemes to articulatory codes: An fMRI study of the role of broca's area in speech production. *Cerebral Cortex*, *19*(9), 2156–2165.

Passingham, R. E. (1996). Attention to action. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *351*(1346), 1473-1479.

Passingham, R. E., & Wise, S. P. (2012). *The neurobiology of the prefrontal cortex: anatomy, evolution, and the origin of insight* (No. 50). Oxford University Press.

Pattamadilok, C., Knierim, I. N., Duncan, K. J. K., & Devlin, J. T. (2010). How does learning to read affect speech perception?. *Journal of Neuroscience*, *30*(25), 8435-8444.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature*, *362*(6418), 342-345.

Paulesu, E., Shallice, T., Danelli, L., & Sberna, M. (2017). *Anatomical Modularity of Verbal Working Memory ? Functional Anatomical Evidence from a Famous Patient with Short-Term Memory Deficits*. *11*(May), 1–16.

Peeva, M. G., Guenther, F. H., Tourville, J. A., Nieto-Castanon, A., Anton, J. L., Nazarian, B., & Alario, F. X. (2010). Distinct representations of phonemes, syllables, and supra-syllabic sequences in the speech production network. *Neuroimage*, *50*(2), 626-638.

Penhune, V. B., & Doyon, J. (2005). Cerebellum and M1 interaction during early learning of timed motor sequences. *NeuroImage*, *26*(3), 801–812.

Peterburs, J., Blevins, L. C., Sheu, Y. S., & Desmond, J. E. (2019). Cerebellar contributions to sequence prediction in verbal working memory. *Brain Structure and Function*, *224*(1), 485–499.

Peterburs, J., Cheng, D. T., & Desmond, J. E. (2016). The association between eye

movements and cerebellar activation in a verbal working memory task. *Cerebral Cortex*, *26*(9), 3802–3813.

Petersen, S. E., Fox, P. T., Posner, M. I., Mintun, M., & Raichle, M. E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, *331*(6157), 585-589.

Petrides, M. (1994). Frontal lobes and behaviour. *Current opinion in neurobiology*, *4*(2), 207-211.

Petrides, M., & Pandya, D. N. (1984). Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *Journal of Comparative neurology*, *228*(1), 105-116.

Petrides, M., & Pandya, D. N. (1999). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *European Journal of Neuroscience*, *16*(2), 291–310.

Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current Opinion in Neurobiology*, *11*(6), 663–672.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as "asymmetric sampling in time." *Speech Communication*, *41*(1), 245–255.

Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The neural correlates of motor skill automaticity. *Journal of Neuroscience*, *25*(22), 5356–5364.

Porrill, J., Dean, P., & Stone, J. V. (2004). Recurrent cerebellar architecture solves the motor-error problem. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1541), 789-796.

Port, R. (2007). How are words stored in memory? Beyond phones and phonemes. *New ideas in psychology*, *25*(2), 143-170.

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., Moore, C. J., Humphreys, G. W., & Wise, R. J. (1997). Segregating semantic from phonological processes during reading. *Journal of cognitive neuroscience*, *9*(6), 727-733.

Pugh, K. R., Shaywitz, B. A., Shaywitz, S. E., Constable, R. T., Skudlarski, P., Fulbright, R. K., & Gore, J. C. (1996). Cerebral organization of component processes in reading. *Brain*, *119*(4), 1221-1238.

Purcell, J. J., Napoliello, E. M., & Eden, G. F. (2011). A combined fMRI study of typed spelling and reading. *NeuroImage*, *55*(2), 750–762.

Rae, C. L., Hughes, L. E., Anderson, M. C., & Rowe, J. B. (2015). The prefrontal cortex achieves inhibitory control by facilitating subcortical motor pathway connectivity. *Journal of Neuroscience*, *35*(2), 786–794.

Raizada, R. D., & Poldrack, R. A. (2007). Selective amplification of stimulus differences during categorical processing of speech. *Neuron*, *56*(4), 726-740.

Rajkowska, G., & Goldman-Rakic, P. S. (1995). DUPLICATE Cytoarchitectonic definition of prefrontal areas in normal human cortex: I. Remapping of areas 9 and 46 and relationship to the Talairach coordinate system. *Cerebral Cortex*, *5*(Table 1), 307–322.

Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nature reviews neuroscience*, *7*(7), 511-522.

Ramnani, N. (2012). Frontal lobe and posterior parietal contributions to the cortico-cerebellar system. *The Cerebellum*, *11*(2), 366-383.

Ramnani, N., & Passingham, R. E. (2001). Changes in the human brain during rhythm learning. *Journal of Cognitive Neuroscience*, *13*(7), 952–966.

Ramnani, N., Toni, I., Josephs, O., Ashburner, J., & Passingham, R. E. (2000). Learning-and expectation-related changes in the human brain during motor learning. *Journal of Neurophysiology*, *84*(6), 3026-3035.

Rapp, B., & Dufor, O. (2011). The neurotopography of written word production: An fMRI investigation of the distribution of sensitivity to length and frequency. *Journal of Cognitive Neuroscience*, *23*(12), 4067–4081.

Rapp, B., & Lipka, K. (2011). The literate brain: The relationship between spelling and reading. *Journal of Cognitive Neuroscience*, *23*(5), 1180–1197.

Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing research*, *271*(1-2), 16-25.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature neuroscience*, *12*(6), 718-724.

Ravizza, S. M., McCormick, C. A., Schlerf, J. E., Justus, T., Ivry, R. B., & Fiez, J. A. (2006). Cerebellar damage produces selective deficits in verbal working memory. *Brain*, *129*(2), 306–320.

Reilly, J. X. O., Beckmann, C. F., Tomassini, V., Ramnani, N., & Johansen-, H.

(2010). *Distinct and Overlapping Functional Zones in the Cerebellum Defined by Resting State Functional Connectivity*. *April*, 953–965.

Reinke, K., Fernandes, M., Schwindt, G., O'Craven, K., & Grady, C. L. (2008). Functional specificity of the visual word form area: General activation for words and symbols but specific network activation for words. *Brain and Language*, *104*(2), 180–189.

Riecker, A., Kassubek, J., Gröschel, K., Grodd, W., & Ackermann, H. (2006). The cerebral control of speech tempo: Opposite relationship between speaking rate and BOLD signal changes at striatal and cerebellar structures. *NeuroImage*, *29*(1), 46-53.

Richter, M., Amunts, K., Mohlberg, H., Bludau, S., Eickhoff, S. B., Zilles, K., & Caspers, S. (2019). Cytoarchitectonic segregation of human posterior intraparietal and adjacent parieto-occipital sulcus and its relation to visuomotor and cognitive functions. *Cerebral Cortex*, *29*(3), 1305–1327.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in neurosciences*, *21*(5), 188-194.

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, *31*(6), 889-901.

Rogalsky, C., Matchin, W., & Hickok, G. (2008). Broca's area, sentence comprehension, and working memory: An fMRI study. *Frontiers in Human Neuroscience*, *2*(OCT), 1–13.

Rosenbaum, D. A. (2009). *Human motor control*. Academic press.

Rottschy, C., Eickhoff, S. B., Schleicher, A., Mohlberg, H., Kujovic, M., Zilles, K., & Amunts, K. (2007). Ventral visual cortex in humans: cytoarchitectonic mapping of two extrastriate areas. *Human brain mapping*, *28*(10), 1045-1059.

Ruan, J., Bludau, S., Palomero-Gallagher, N., Caspers, S., Mohlberg, H., Eickhoff, S. B., Seitz, R. J., & Amunts, K. (2019). *Probabilistic cytoarchitectonic map of Area 6ma (preSMA, mesial SFG) (v9.1)* [Data set]. Human Brain Project Neuroinformatics Platform.

Roy, C. S., & Sherrington, C. S. (1890). On the regulation of the blood-supply of the brain. *The Journal of physiology*, *11*(1-2), 85-158.

Rushworth, M. F. S., Behrens, T. E. J., & Johansen-Berg, H. (2006). Connection patterns distinguish 3 regions of human parietal cortex. *Cerebral Cortex*, *16*(10), 1418–1430.

Rushworth, M. F. S., Krams, M., & Passingham, R. E. (2001). The attentional role of the left parietal cortex: The distinct lateralization and localization of motor attention in the human brain. *Journal of Cognitive Neuroscience*, *13*(5), 698–710.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and Learning) out of the Corner of Your Ear. *Psychological Science*, *8*(2), 101–105.

Sakai, K., Rowe, J. B., & Passingham, R. E. (2002a). *Parahippocampal Reactivation Signal at Retrieval after Interruption of Rehearsal*. *22*(15), 6315–6320.

Sakai, K., Ramnani, N., & Passingham, R. E. (2002b). Learning of sequences of finger movements and timing: frontal lobe and action-oriented representation. *Journal of neurophysiology*, *88*(4), 2035-2046.

Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of verbal learning and verbal behavior*, *21*(2), 150-164.

Salin, P. A., Malenka, R. C., & Nicoll, R. A. (1996). Cyclic AMP mediates a presynaptic form of LTP at cerebellar parallel fiber synapses. *Neuron*, *16*(4), 797-803.

Salmon, E., Van Der Linden, M., Collette, F., Delfiore, G., Maquet, P., Degueldre, C., Luxen, A., & Franck, G. (1996). Regional brain activity during working memory tasks. *Brain*, *119*(5), 1617–1625. h

Saura, D., Kreher, B. W., Schnell, S., Kümmerera, D., Kellmeyera, P., Vrya, M. S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(46), 18035–18040.

Savin, H. B., & Bever, T. G. (1970). The nonperceptual reality of the phoneme. *Journal of Verbal Learning and Verbal Behavior*, *9*(3), 295–302.

Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension—an fMRI study. *Cerebral Cortex*, *22*(7), 1662-1670.

Scheperjans, F., Hermann, K., Eickhoff, S. B., Amunts, K., Schleicher, A., & Zilles, K. (2007). Observer-independent cytoarchitectonic mapping of the human superior parietal cortex. *Cerebral Cortex*, *18*(4), 846-867.

Schmahmann, J. D. (1996). From movement to thought: anatomic substrates of the cerebellar contribution to cognitive processing. *Human brain mapping*, *4*(3), 174-198.

Schmahmann, J. D., & Pandya, D. N. (1997a). Anatomic organization of the basilar pontine projections from prefrontal cortices in rhesus monkey. *Journal of Neuroscience*, *17*(1), 438-458.

Schmahmann, J. D., & Pandya, D. N. (1997b). The cerebrocerebellar system. *International review of neurobiology*, *41*, 31-60.

Schmahmann, J. D., Doyon, J., McDonald, D., Holmes, C., Lavoie, K., Hurwitz, A. S., & Petrides, M. (1999). Three-dimensional MRI atlas of the human cerebellum in proportional stereotaxic space. *Neuroimage*, *10*(3), 233-260.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1–2), 1–46.

Shallice, T., & Butterworth, B. (1977). Short-term memory impairment and spontaneous speech. *Neuropsychologia*, *15*(6), 729-735.

Shallice, T., & Papagno, C. (2019). Impairments of auditory-verbal short-term memory: Do selective deficits of the input phonological buffer exist? *Cortex*, *112*, 107-121.

Sharp, D. J., Bonnelle, V., De Boissezon, X., Beckmann, C. F., James, S. G., Patel, M. C., & Mehta, M. A. (2010). Distinct frontal systems for response inhibition, attentional capture, and error processing. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(13), 6106–6111.

Shuster, L. I., & Lemieux, S. K. (2005). An fMRI investigation of covertly and overtly produced mono- and multisyllabic words. *Brain and Language*, *93*(1), 20–31.

Sigl, B., Caspers, S., Bludau, S., Mohlberg, H., Eickhoff, S. B., & Amunts, K. (2019). Probabilistic cytoarchitectonic map of Area 6d3 (SFS) (v4.1) [Data set]. Human Brain Project Neuroinformatics Platform.

Silveri, M. C., Di Betta, A. M., Filippini, V., Leggio, M. G., & Molinari, M. (1998). Verbal short-term store-rehearsal system and the cerebellum. Evidence from a patient with a right cerebellar lesion. *Brain*, *121*(11), 2175–2187.

Sjöblom, A., (2019) Embodied word-form learning: the motoric and perceptual determinants of verbal sequence learning. (Doctoral dissertation, Royal Holloway, University of London).

Sjöblom, A., & Hughes, R. W. (2020). The articulatory determinants of verbal sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Smith, E. E., Jonides, J., & Koeppe, R. A. (1996). Dissociating verbal and spatial working memory using PET. *Cerebral cortex*, *6*(1), 11-20.

St-Louis, M. È., Hughes, R. W., Saint-Aubin, J., & Tremblay, S. (2019). The resilience of verbal sequence learning: Evidence from the hebb repetition effect. *Journal of Experimental Psychology: Learning Memory and Cognition*, *45*(1), 17–25.

Stadler, M. A. (1993). Implicit serial learning: Questions inspired by Hebb (1961). *Memory & cognition*, *21*(6), 819-827.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652–654.

Sternberg, S., Wright, C. E., Knoll, R. L., & Monsell, S. (1980). Motor programs in rapid speech: Additional evidence. In *Perception and Production of Fluent Speech* (pp. 507–534).

Stokes, M., Saraiva, A., Rohenkohl, G., & Nobre, A. C. (2011). Imagery for shapes activates position-invariant representations in human visual cortex. *NeuroImage*, *56*(3), 1540–1545.

Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience*, *29*(5), 1565–1572.

Stoodley, C. J., & Schmahmann, J. D. (2009). Functional topography in the human cerebellum: a meta-analysis of neuroimaging studies. *Neuroimage*, *44*(2), 489-501.

Stoodley, C. J., & Schmahmann, J. D. (2010). Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex*, *46*(7), 831–844.

Szmalec, A., Duyck, W., Vandierendonck, A., Mata, A. B., & Page, M. P. A. (2009). The Hebb repetition effect as a laboratory analogue of novel word learning. *Quarterly Journal of Experimental Psychology*, *62*(3), 435–443.

Szmalec, A., Loncke, M., Page, M., & Duyck, W. (2011). Order or disorder? Impaired Hebb learning in dyslexia. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1270.

Szmalec, A., Page, M. P., & Duyck, W. (2012). The development of long-term lexical representations through Hebb repetition learning. *Journal of Memory and Language*, *67*(3), 342-354.

Talairach, J., & Tournoux, P. (1988). Co-planar stereotaxic atlas of the human brain. *Theime, Stuttgart, Ger*, *270*(132), 90128-5.

Tan, L., & Ward, G. (2002). A recency-based account of the list length effect in free recall. *Memory and Cognition*, *30*(6), 885–892.

Tanji, J., & Hoshi, E. (2001). Behavioral planning in the prefrontal cortex. *Current Opinion in Neurobiology*, *11*(2), 164–170.

Taylor, J. C., Macken, B., & Jones, D. M. (2015). A matter of emphasis: Linguistic stress habits modulate serial recall. *Memory and Cognition*, *43*(3), 520–537.

Thach, W. T. (1998). A role for the cerebellum in learning movement coordination. *Neurobiology of learning and memory*, *70*(1-2), 177-188.

Thach, W. T., Goodkin, H. P., & Keating, J. G. (1992). The cerebellum and the adaptive coordination of movement. *Annual Review of Neuroscience*, *15*, 403–442.

Tomlinson, S. P., Davis, N. J., Morgan, H. M., & Bracewell, R. M. (2014). *Cerebellar Contributions to Verbal Working Memory*. 354–361.

Toni, I., Thoenissen, D., & Zilles, K. (2001). Movement preparation and motor intention. *Neuroimage*, *14*(1), S110-S117.

Tree, J. J., Longmore, C., & Besner, D. (2011). Orthography, phonology, short-term memory and the effects of concurrent articulation on rhyme and homophony judgements. *Acta psychologica*, *136*(1), 11-19.

Turner, B. H., Mishkin, M., & Knapp, M. (1980). Organization of the amygdalopetal projections from modality-specific cortical association areas in the monkey. *Journal of Comparative Neurology*, *191*(4), 515-543.

Tzvi, E., Münte, T. F., & Krämer, U. M. (2014). Delineating the cortico-striatal-cerebellar network in implicit motor sequence learning. *Neuroimage*, *94*, 222-230.

Uddin, L. Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D. A., Greicius, M. D., & Menon, V. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus: Evidence from functional and structural connectivity. *Cerebral Cortex*, *20*(11), 2636–2646.

Ugolini, G., & Kuypers, H. G. J. M. (1986). Collaterals of corticospinal and pyramidal fibres to the pontine grey demonstrated by a new application of the fluorescent fibre labelling technique. *Brain research*, *365*(2), 211-227.

Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, *72*(2), 385–396.

Van der Heijden, A. H. C. (1993). The role of position in object selection in vision. *Psychological Research, 56*(1), 44-58.

Vallar, G., & Baddeley, A. D. (1984). Phonological short-term store, phonological processing and sentence comprehension: A neuropsychological case study. *Cognitive Neuropsychology*, *1*(2), 121–141.

Vallar, G., & Baddeley, A. (1987). Phonological short-term store and sentence processing. *Cognitive Neuropsychology*, *4*(4), 417-438.

Vallar, G., & Papagno, C. (1995). Neuropsychological impairments of short-term memory.

Van der Heijden, A. H. (2003). *Selective attention in vision*. Routledge.

Voogd, J., & Glickstein, M. (1998). The anatomy of the cerebellum. *Trends in cognitive sciences*, *2*(9), 307-313.

Wakana, S., Jiang, H., Nagae-Poetscher, L. M., Van Zijl, P. C. M., & Mori, S. (2004). Fiber Tract-based Atlas of Human White Matter Anatomy. *Radiology*, *230*(1), 77–87.

Wandell, B. A. (2011). The neurobiological basis of seeing words. *Annals of the New York Academy of Sciences*, *1224*(1), 63.

Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory: Functional equivalence across modalities. *Memory*, *13*(3–4), 308–317.

Warrington, E. K., & Shallice, T. (1969). The selective impairment of auditory verbal short-term memory. *Brain*, *92*(4), 885-896.

Watkins, M. J., Watkins, O. C., & Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior*, *13*(4), 430–447.

Weinrich, M., & Wise, S. P. (1982). The premotor cortex of the monkey. *Journal of Neuroscience*, *2*(9), 1329-1345.

Wildgruber, D., Kischka, U., Ackermann, H., Klose, U., & Grodd, W. (1999). Dynamic pattern of brain activation during sequencing of word strings evaluated by fMRI. *Cognitive Brain Research*, *7*(3), 285-294.

Willingham, D. B. (1998). A Neuropsychological Theory of Motor Skill Learning. *Psychological Review*, *105*(3), 558–584.

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin and Review*, *8*(1), 44–57.

Wilson, M., & Emmorey, K. (1997). A visuospatial "phonological loop" in working memory: Evidence from American Sign Language. *Memory and Cognition*, *25*(3), 313–320.

Wilson, M., & Emmorey, K. (1998). A "word length effect" for sign language: Further evidence for the role of language in structuring working memory. *Memory and Cognition*, *26*(3), 584–590.

Wilson, M., & Fox, G. (2007). Working memory for language is not special: Evidence for an articulatory loop for novel stimuli. *Psychonomic Bulletin and Review*, *14*(3), 470–473.

Wise, S. P. (1985). The primate premotor cortex: past, present, and preparatory. *Annual review of neuroscience*, *8*(1), 1-19.

Wise, R. J. S., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within "Wernicke's area." *Brain*, *124*(1), 83–95.

Wolpert, D M, & Kawato, M. (1998). *Multiple paired forward and inverse models for motor control*. *11*, 1317–1329.

Wolpert, D M, & Miall, R. C. (1996). Forward Models for Physiological Motor

Control. *Neural Networks*, *9*(8), 1265–1279.

Wolpert, Daniel M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–347.

Woodward, A. J., Macken, W. J., & Jones, D. M. (2008). Linguistic familiarity in short-term memory: A role for (co-)articulatory fluency? *Journal of Memory and Language*, *58*(1), 48–65.

Worsley, K. J., Marrett, S., Neelin, P., Vandal, A. C., Friston, K. J., & Evans, A. C. (1996). A unified statistical approach for determining significant signals in images of cerebral activation. *Human brain mapping*, *4*(1), 58-73.

Wray, A. (2015). Why are we so sure we know what a word is?. in *The Oxford Handbook of the Word*, ed. J. Taylor (Oxford: Oxford University Press).

Xue, G., Aron, A. R., & Poldrack, R. A. (2008). Common neural substrates for inhibition of spoken and manual responses. *Cerebral Cortex*, *18*(8), 1923-1932.

Yanaoka, K., Nakayama, M., Jarrold, C., & Saito, S. (2019). Determining the developmental requirements for hebb repetition learning in young children: Grouping, short-term memory, and their interaction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(4), 573.

Yarkoni, T., Speer, N. K., Balota, D. A., McAvoy, M. P., & Zacks, J. M. (2008). Pictures of a thousand words: Investigating the neural mechanisms of reading with extremely rapid event-related fMRI. *NeuroImage*, *42*(2), 973–987.

Yeatman, J. D., Rauschecker, A. M., & Wandell, B. A. (2013). Anatomy of the visual word form area: Adjacent cortical circuits and long-range white matter connections. *Brain and Language*, *125*(2), 146–155.

Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*(5058), 846-849.

Zevin, J. D., & McCandliss, B. D. (2005). Dishabituation of the BOLD response to speech sounds. *Behavioral and Brain Functions*, *1*(1), 1-12.

APPENDICES

Appendix A

Pilot Experiments

Visual stimuli were used for the first and second pilot experiments prior to the pilot experiments conducted with auditory stimuli. For the experiments proper, however, the auditory study (Chapter III) is reported before the visual study (Chapter IV).

6.1 Pilot 1 (using visual stimuli)

6.1.1 Subjects

Eight students (*female*= 5, *male* = 3, *mean age* = 26 years, *SD*= 3.54) were recruited from Royal Holloway, University of London's Experiment management system and remunerated in course credits for their participation. All subjects were right-handed and reported being psychologically and neurologically healthy. Two of these subjects were monolingual native British English speakers, two were non-native British English speakers and two were native British English speakers but also bilingual. No subjects reported any history of language, speech, memory, auditory or visual disorders.

6.1.2 Results

6.1.2.1 Experiment A.

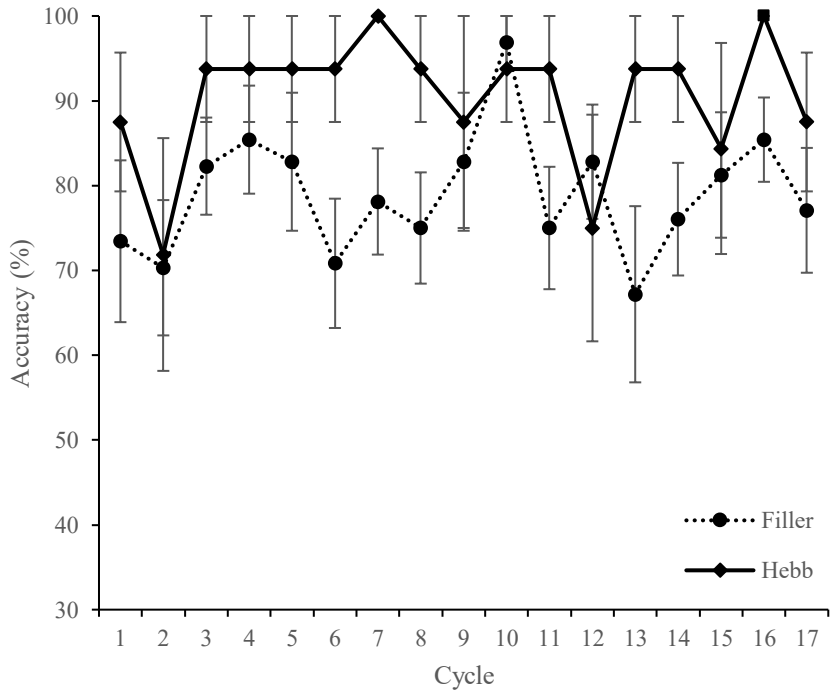
A one tailed t-test was conducted to compare recall accuracy for similar lists (*M*= 59.64%, *SE*= 3.32) to dissimilar lists (*M*= 70.05%, *SE*= 3.66). Results of the one-tailed t-test showed that a significant difference was found $t(7) = -1.86$, $p = .05$.

6.1.2.2 Experiment B.

The recall accuracy across the 17 cycles, for both Filler and Hebb sequences, is shown in Figure 24. A 2 (list-type) × 17 (cycle) repeated measures ANOVA was conducted to assess performance accuracy for Filler (*M*= 78.98%, *SE*= 4.17) and Hebb (*M*= 90.44, *SE*= 2.57) trials across the cycles. There was a reliable main effect of list-type $F(1,7) = 11.23$, $MSE = 795.22$, $p = .012$, $\eta_p^2 = .616$. The non-significant result of cycle $F(16,112) = 1.041$, $MSE = 493.6$, $p = .421$, $\eta_p^2 = .129$ is likely related to the fact that the Hebb sequence was rapidly learned. The list-type × cycle interaction was also non-significant, $F(16,112) = .895$, $MSE = 401.6$, $p = .576$, $\eta_p^2 = .113$.

Figure 24

Serial recall accuracy for Filler and Hebb trials across the 17 cycles in Pilot 1 (visual).



6.2 Pilot 2 (using visual stimuli)

6.2.1 Design and Procedure

Performance in Pilot 1 showed a ceiling effect early on in sub-experiment B (cycle 7) and the full extent of learning could therefore not be assessed accurately. It is possible that a ceiling effect with phonologically dissimilar letters precluded the capacity to observe a difference between performance in Hebb and Filler trials. To combat this, three aspects of the design were changed. Firstly, the first letter of any sequence would not appear in the sequence fragment at the recall stage, meaning that the first letter was always removed and included in the choice of four letters to place back into the sequence. This was to prevent the first letter acting as a strong cue for the rest of the sequence, especially in relation to the Hebb sequence. Secondly, phonologically similar letters were now used for Filler and Hebb trials instead of phonologically dissimilar letters as used in Pilot 1 sub-experiment B. This was in an attempt to increase task difficultly and reduce the behavioural accuracy for both Filler and Hebb trials at the start of the experiment, thus enabling observation of a steadier learning curve for Hebb trials. Thirdly, the order of the trial manipulations across the sub-experiment was changed. Sub-experiment B in pilot 1 began with a Hebb trial. Two subsequent Filler trials separated the next occurrence of a Hebb trial. Instead, in Pilot 2 sub-experiment B, the first Hebb trial occurred after three Filler trials at the start of the experiment. Changing the spacing of Hebb trials. It is worth noting that although the visual experiments are presented second to the auditory experiments in the main thesis, the visual experiments were actually conducted first.

6.2.2 Subjects

Seven students (*female*= 6, *male*=1 *mean age* = 20 years, *SD*= 2.16) were recruited from Royal Holloway, University of London's Experiment management system and reimbursed in credits for their participation. All subjects were right-handed and psychologically and neurologically healthy. Five of these subjects were monolingual native British English speakers, one was native British English speaker but bilingual and two were non-native English speakers. No subjects reported a history of language, speech, memory, auditory or visual disorders.

6.2.3 Results

6.2.3.1 Experiment A.

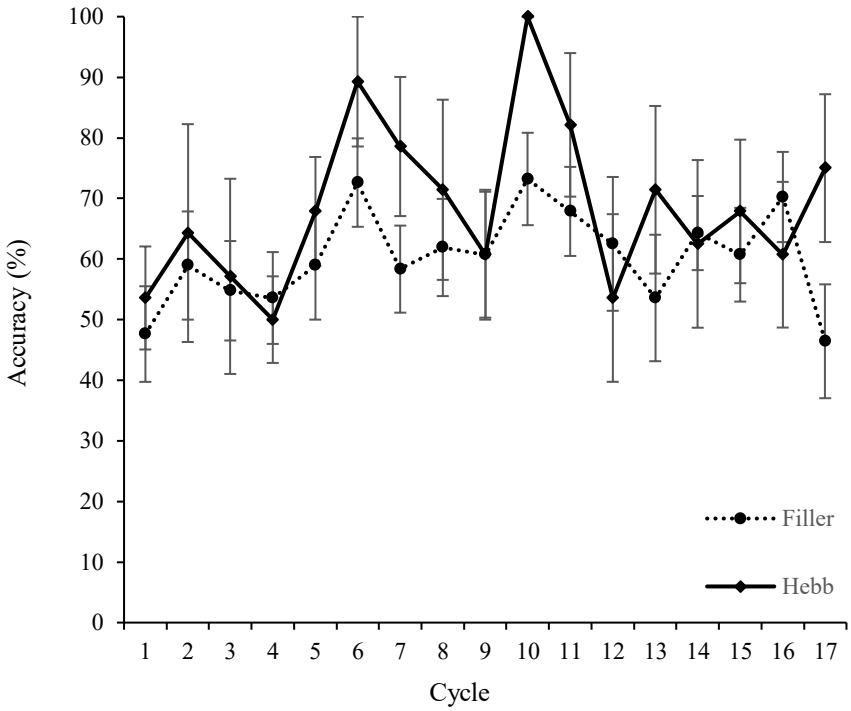
Whilst there was again a numerical trend for a phonological similarity effect where similar lists (*M*= 61.9%, *SE*= 3.65) were more poorly recalled than dissimilar lists (*M*= 67.56%, *SE*= 3.88) the difference was not significant on this occasion $t(7) = -1.03$, $p = .34$.

6.2.3.2 Experiment B.

Figure 25 shows Filler (*M*= 60.22%, *SE*= 4.5) and Hebb (*M*= 70.8%, *SE*= 5.5) accuracy. It is evident that changing the letter set to phonologically similar had an effect on performance accuracy, as regardless of condition, performance accuracy began at around 50% from the outset rather than above 70% in pilot 1 where phonologically dissimilar letters were used. A 2 (list-type) x 17 (cycle) repeated measures ANOVA showed a main effect of cycle, $F(16,96) = 1.952$, $MSE = 631.653$, $p = .024$, $\eta_p^2 = .245$ but list-type $F(1, 6) = 3.896$, $MSE = 1707.6.653$, $p = .096$, $\eta_p^2 = .394$ and the list-type by cycle interaction $F(1, 16) = .453$, $MSE = 340.949$, $p = .963$, $\eta_p^2 = .07$, were both non-significant.

Figure 25

Recall accuracy for Filler and Hebb trials across the 17 cycles in the second pilot for Experiment 1.



6.3 Stimulus identification task for auditory pilot experiments

The stimulus identification task was conducted to assess the intelligibility of the spoken recordings when presented concurrently with the simulated noise of an MRI scanner. Subjects were required to press the keyboard letter corresponding to what they heard over the headphones. The keys on the keyboard corresponding to the 14 letters in question were each marked with a sticker. Subjects completed the task sat at a desk wearing in-ear Apple earphones connected to a computer playing the stimuli as well as over-ear Sennheiser HD 380 pro headphones playing scanner noise. Volume level was tracked and adjusted for subjects upon request. Eight of the ten subjects were native monolingual English speakers and two were non-native English speakers, none

of whom took part in the fMRI experiments. The results showed that 9 subjects scored 100% accuracy in recognising the stimuli while one subject scored 90% but this, according to a spontaneous post-task report by the subject, was due to an accidental pressing of the key adjacent to the correct one. It was concluded, therefore, that all the sound stimuli were clearly intelligible and that they would very likely be so in the MRI scanner too.

6.4 Pilot 3 (using auditory stimuli)

6.4.1 Subjects

Six students (all female, *mean age* = 21. 5 years, *SD* = 3.2) were recruited from Royal Holloway, University of London's Experiment management system and remunerated in course credits for their participation. This pilot was conducted using the same behavioural set up and the mock MRI scanner as the pilot experiments during Experiment 1. All subjects were right-handed, psychologically and neurologically healthy. Four of these subjects were monolingual native British English speakers and two were non-native English speakers. All subjects reported no history of language, speech, memory, auditory or visual disorders.

6.4.2 Results

6.4.2.1 Experiment A.

A paired t-test was conducted to compare performance accuracy of similar (*M*= 51.04%, *SE*= 3.9) lists to dissimilar lists (*M*= 52.43%, *SE*= 4.0), $t(5) = -0.34$, $p = .74$. The difference was not significant on this occasion.

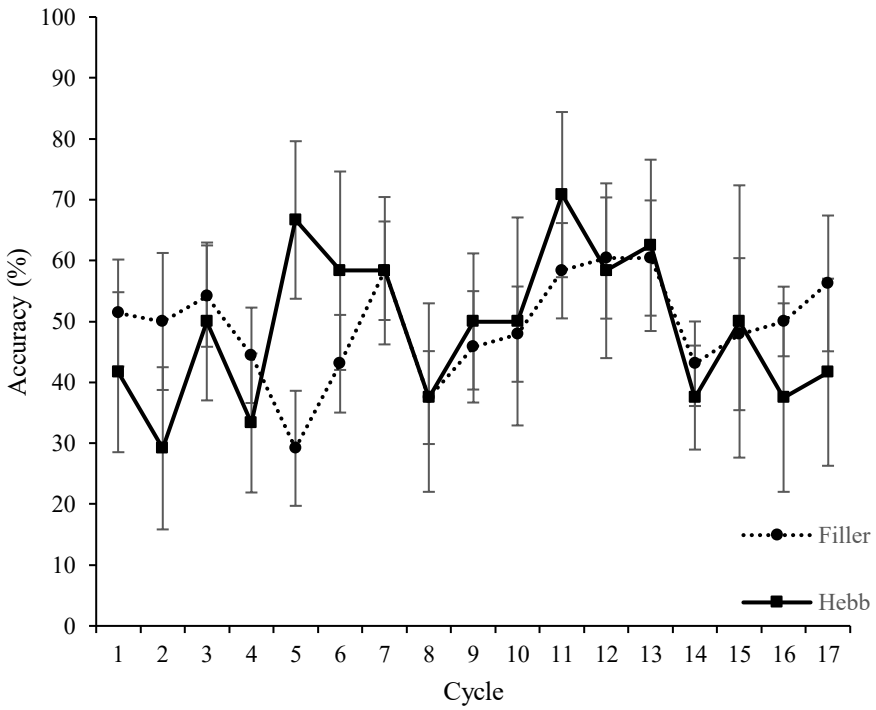
6.4.2.2 Experiment B.

As shown in Figure 26 the data from Experiment B contains a considerable amount of noise and the lack of significant results are likely attributable to the low number of

participants and the technical issues encountered with the button box subjects used to complete the task. A numerical difference between Filler (*M*= 48.73%, *SE*= 6.5) and Hebb (*M*= 51%, *SE*= 6.12) sequences was observed. A 2 (list-type) × 17 (cycle) repeated measures ANOVA was conducted to assess the difference between recall of Filler and Hebb lists but showed no main effects; list-type, $F(1, 5) = 1.65$, $MSE = 1563.606$, $p = .702$, $\eta_p^2 = .032$, cycle, $F(16,80) = 1.018$, $MSE = 760.759$, $p = .447$, $\eta_p^2 = .169$, and list-type by cycle interaction , $F(16,80) = .878$, $MSE = 778.197$, $p = .596$, $\eta_p^2 = .149$.

Figure 26

Graph showing group accuracy across 17 cycles for Filler and Hebb trials, trendlines included.



6.5 Pilot 4 (using auditory stimuli)

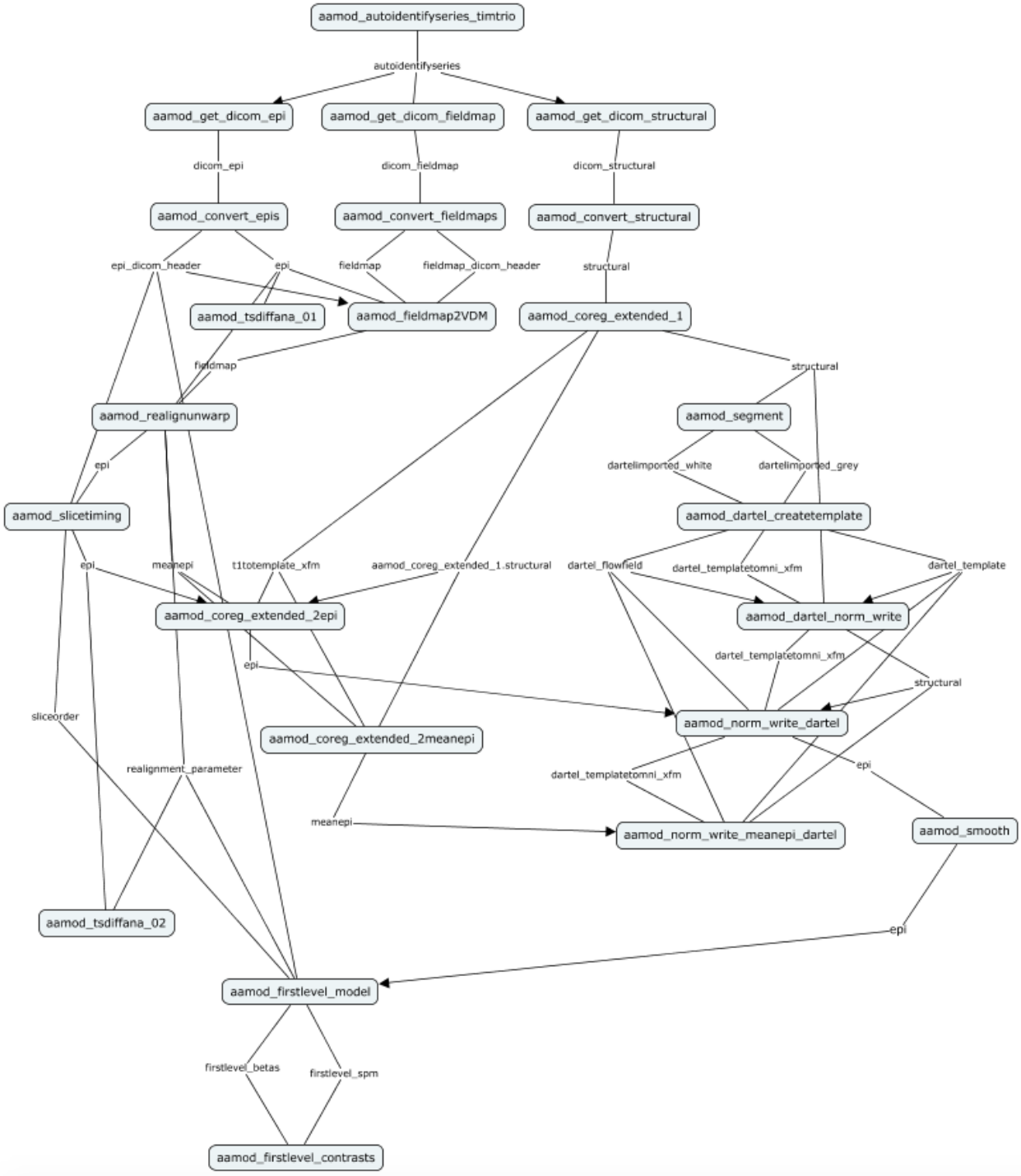
Due to technical issues during the first pilot experiment, the second was conducted to ensure that the MATLAB code worked with a changed hardware setup using a NATA button box and National Instruments Board (see Chapter 2: Section 2.6.1, Figure 6). Pilot 4 was conducted from the MRI control room in the CUBIC facility with one subject. The subject was a right-handed, native monolingual speaker of English reporting no history of language, speech, memory, auditory, visual, neurological or psychological disorders

Appendix B

Conceptual Map

Figure 27

Conceptual map illustrating order of pre-processing steps (generated by aa-pipeline report module).

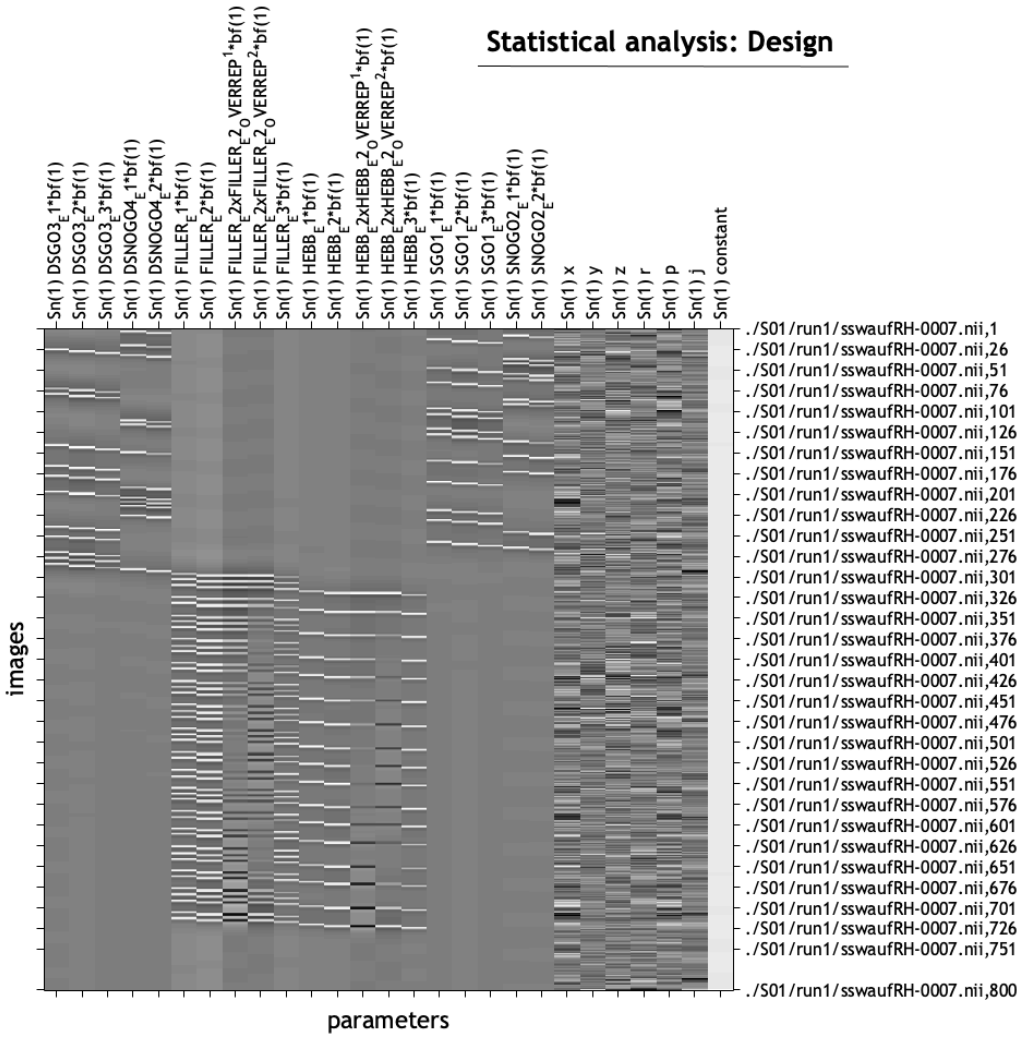


Appendix C

First-level model

Figure 28

Image of first-level model taken from SPM.

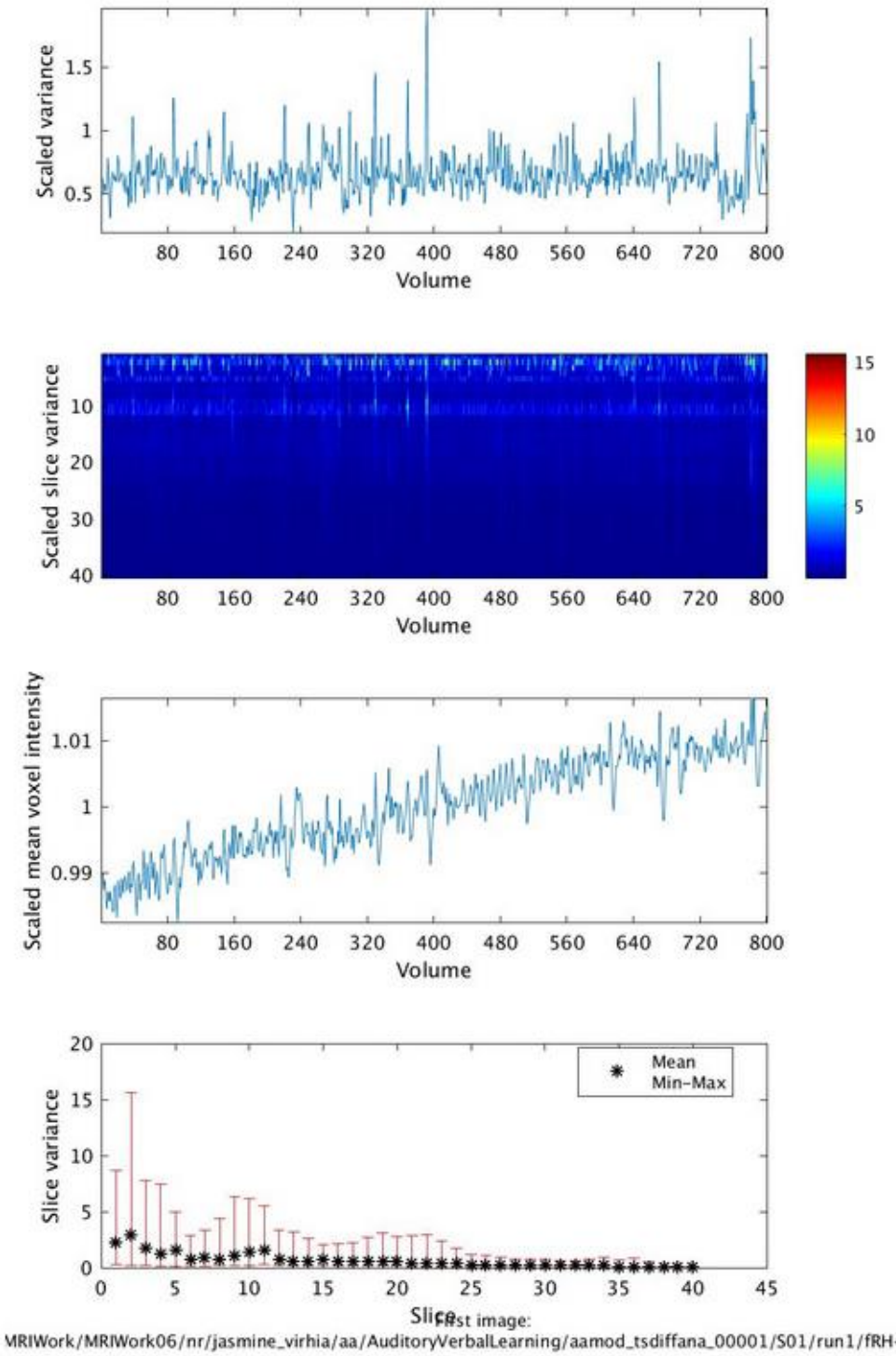


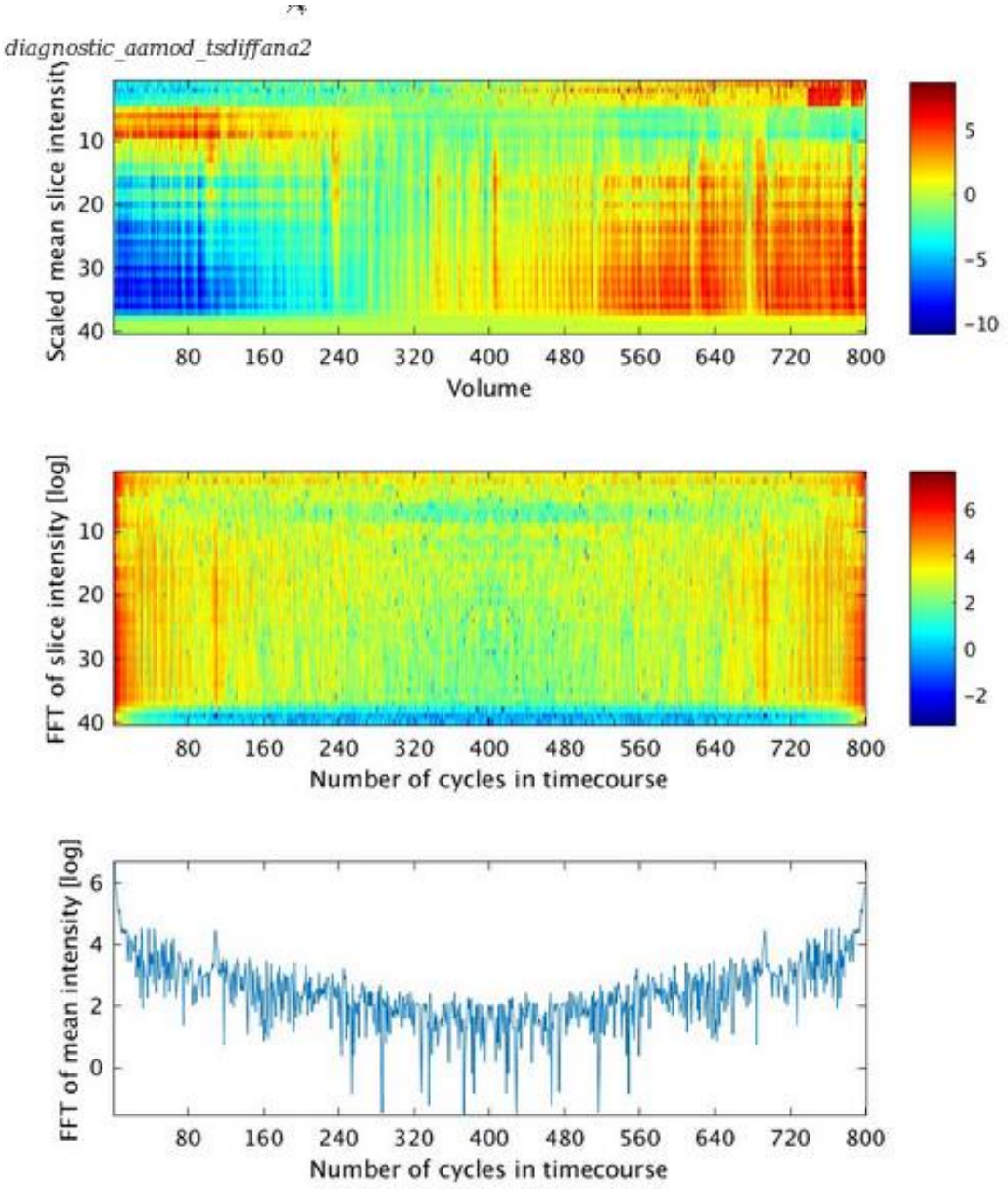
Appendix D

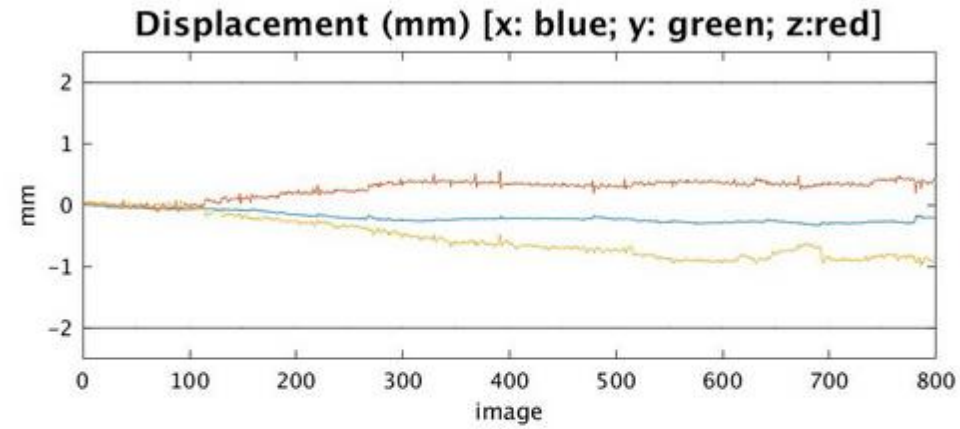
fMRI Data Quality Assurance diagnostic graphs

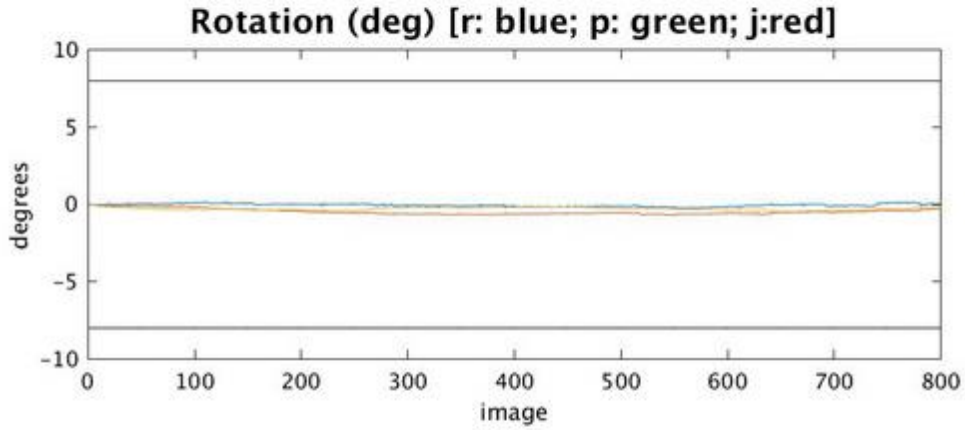
Figure 29

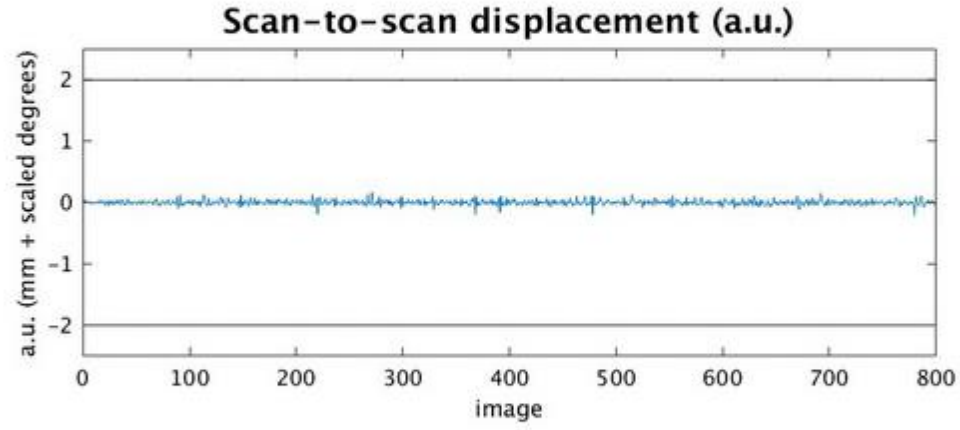
Graphs showing plots of; scaled variance across volumes for the whole volume (first plot) and for each slice (second plot); scaled mean voxel intensity across volumes for the whole volume (third plot) and for each slice (fifth plot); the maximum, mean, and minimum within-slice variance for each slice (fourth plot); and the log transformed Fast-Fourier Transform (FFT) of slice (sixth plot) and volume (seventh plot) intensity.




diagnostic_aamod_tsdiffana2
Scaled mean slice intensity
10
20
30
40
80
160
240
320
400
480
560
640
720
800
Volume
5
0
-5
-10
FFT of slice intensity [log]
10
20
30
40
80
160
240
320
400
480
560
640
720
800
Number of cycles in timecourse
6
4
2
0
-2
FFT of mean intensity [log]
6
4
2
0
80
160
240
320
400
480
560
640
720
800
Number of cycles in timecourse







Movement maximums

Sess	x	y	z	rotx	roty	rotz
1	0.007	0.549	0.073	0.003	0.000	0.000

Appendix E

Boxplots showing outliers in MRI Data based on Motion Correction Parameters

Figure 30

Boxplot showing outlier subjects based on motion correction parameters for MRI Experiment 1 (auditory).

Outlier(s) in Trans - x: None

Outlier(s) in Trans - y: None

Outlier(s) in Trans - z: 18

Outlier(s) in Pitch: None

Outlier(s) in Roll: None

Outlier(s) in Yaw: None

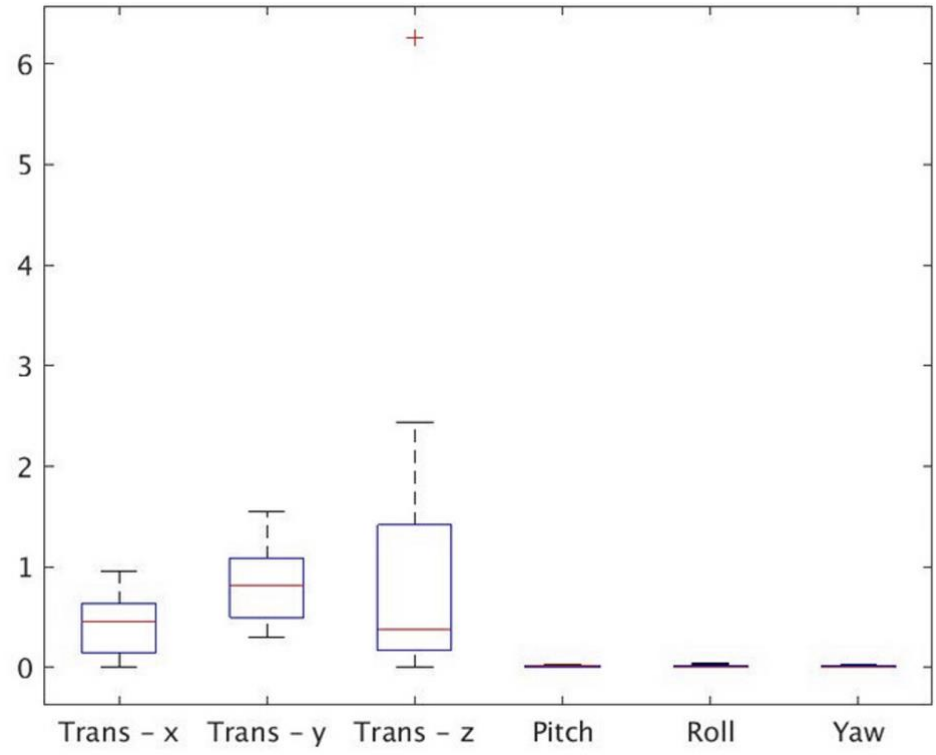


Figure 31

Boxplot showing outlier subjects based on motion correction parameters for MRI Experiment 2 (visual).

Outlier(s) in Trans - x: None

Outlier(s) in Trans - y: None

Outlier(s) in Trans - z: 18

Outlier(s) in Pitch: None

Outlier(s) in Roll: 9 11

Outlier(s) in Yaw: 2 11